

# Multi-Resolution Satellite Data Processing for Contextually Relevant Land Use–Land Cover Mapping

## Abstract

Scrub-shrub areas and rain-fed farmlands form a major share of India's semi-arid and dryland landscapes but are difficult to distinguish reliably in global land-use products derived from satellite data, such as Google's Dynamic World, because these classes have highly similar spectral signatures. The absence of large, geographically distributed, high-quality training datasets prevents current state-of-the-art (SOTA) global land-cover models from accurately capturing the differences and restricts their usefulness for restoration planning, agroforestry, and ecological monitoring. To address this challenge, we present a scalable pipeline that automatically identifies a diverse samples of high-res (1.19 m) RGB satellite map tiles from across India, and applies computer vision methods on these tiles to accurately delineate the two classes, then uses the delineated areas to automatically create training datasets that can be used on med-res (10m) spectral satellite data to produce wall-to-wall maps. The sampling is done using a greedy information theoretic algorithm that can be applied on a less accurate land-use product, to eventually build improved classifiers that can transform it into a more accurate end product. This strategy keeps monetary and compute costs low by having to acquire and process hi-res imagery for only 3% of the geographic area. The methodology is fairly generic and can be adapted to improving the classification of other land-use classes too. In addition to successfully improving the accuracy of shrub-scrub and farm classes, we also show that a new class of agro-horticulture plantations can be introduced in a similar manner. Finally, we use this method to produce pan-India annual datasets of land-use from 2017 to 2025.

## 1 Introduction

### 1.1 Background and Motivation

Scrublands play a vital role in India's ecological and socioeconomic fabric. [18] [19] They act as transitional ecosystems between forests and croplands, supporting biodiversity, groundwater recharge, and pastoral livelihoods in semi-arid and dryland regions. For millions of rural and nomadic communities, these landscapes are not wastelands but critical commons that provide fodder, fuelwood, and grazing support. [19]

However, due to its misclassification as wasteland in administrative framework under historical utilitarian logic[15][14][7], these scrublands are often neglected in restorative programs [8] [18] or are prone to degradation, agricultural conversions, urban encroachments, infrastructure development and tree plantations[15] [1]. This degradation is often overlooked in national land management programs because scrublands are poorly mapped within existing Land Use Land Cover (LULC) [28] [17] [18] [2] [9] [30].

Accurate delineation of scrublands is thus essential for multiple domains: from ecological restoration and pasture management to climate resilience planning and rural livelihood policy. However, despite growing remote sensing capabilities around the world,

the accurate distinction between scrublands and farms remains a persistent challenge in mapping efforts in India [28] [17] [18].

Agro-plantations, on the otherhand, represent a rapidly expanding land-use category across India, ranging from state-supported horticultural plantations to large private agro-forestry blocks that now span millions of hectares nationwide [22]. Despite this substantial footprint, plantations are typically subsumed under broad "tree cover" or "agricultural land" classes in most national and global LULC products. [2] [9] [30] This misrepresentation affects not just statistical accuracy but also policy interpretations, for example, when estimating deforestation rates, evaluating afforestation drives or even evaluating true change in cropping intensity. [5] The lack of high-quality, geographically diverse training samples makes plantations a favourable category for demonstrating how our pipeline can be extended to incorporate new important land-use classes with minimal manual effort.

### 1.2 Problem Definition

The core challenge in distinguishing scrublands from rain-fed farmlands arises from their inherent spectral and temporal similarity. Across much of semi-arid India, both land-use types are strongly monsoon-dependent and exhibit comparable seasonal greening patterns. During peak vegetation periods, spectral indices such as NDVI and EVI show nearly identical reflectance trajectories for the two classes. Consequently, pixel-based classifiers—especially those trained on globally aggregated datasets—struggle to capture the subtle, locally driven nuances required to reliably separate them.

A fundamental driver of this confusion is the lack of large, high-quality, geographically diverse labeled datasets focused specifically on scrubland and rain-fed farmland. Without representative samples from India's varied agro-ecological zones (AEZs), models tend to overfit to limited regional characteristics and fail to generalize nationally. SOTA global land-cover products such as Google's Dynamic World[2], ESA WorldCover[30], and ESRI Land Cover[9] offer valuable global consistency but perform poorly in the Indian context. Their global training data, does not adequately capture India's monsoon-dependent, smallholder-dominated landscapes. As a result, these models frequently misclassify scrublands as croplands and vice versa, particularly in low-input, rain-fed regions.

However, assembling such a dataset manually is extremely difficult and time-consuming. India's vast geographic extent makes ground-based or manual annotation approaches prohibitively expensive and operationally infeasible. Unlike many Western regions, where well-curated, publicly available land-use datasets and field inventories support large-scale model training, India suffers from limited labeled data, inconsistent field records, and sparse region-specific mapping efforts. Creating the hundreds of thousands of high-quality samples required to teach medium-resolution models to reliably separate these classes would demand substantial expert labor and would be practically impossible to scale across all AEZs.

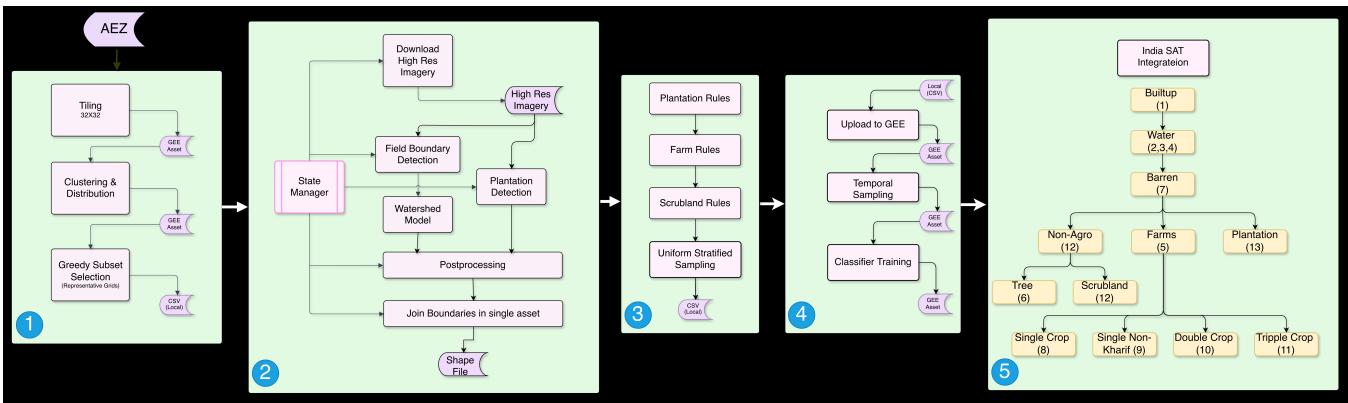


Figure 1: This figure shows the complete pipeline. 1. Spatial Representativeness and Tile Selection, 2. High-Resolution Boundary Delineation, 3. Entropy and Geometry Based Boundary Refinement 4. Sample Generation and Classifier Training, 5. Integration into the existing LULC Framework

### 1.3 Present Study and Contributions

To address these challenges, we propose a methodology that uses ground truth derived through computer vision methods on hi-res data, as opposed to groundtruth which is normally marked through desk work or field surveys. We presents a viable method to create a large volume of groundtruth from limited hi-res data to train satellite data models that can capture fine-scale spectral, temporal, and morphological nuances which are unique to each agro-ecological zones.

The key contributions of this work are as follows:

- (1) **Sampling strategy for selecting hi-res labeling regions:**  
We propose a sampling method to identify where hi-resolution ground truth should be constructed. By operating on medium-resolution feature embedding clusters and explicitly promoting diversity across all agro-ecological zones, this method yield a set of tiles that are representative and information-rich.
- (2) **Framework for CV-derived ground truth at scale** We design a generalizable framework that leverages high-resolution imagery and computer-vision pipelines to automatically derive ground-truth labels for distinct land-use types such as scrublands, farms, and plantations. The framework enables the generation of high-quality samples across India without requiring exhaustive manual annotation, providing a scalable and cost-effective alternative for regions that are typically underrepresented due to limited availability of field surveys or desk-based mapping efforts.
- (3) **Nationwide annual LULC layers** As an outcome of this methodology, we generated annual LULC layers from 2017-2025 with improved scrubland and farm distinction along with novel agro-plantation class.

Through these contributions, this study highlights how targeted high-resolution sampling and localized model training can bridge the accuracy gap in global models, paving the way for a new generation of regionally contextualized LULC products.

## 2 Related Work

**Global 10 m land-cover products.** Recent global LULC products such as Dynamic World, Esri Global Land Cover, and ESA WorldCover provide 10 m, wall-to-wall maps from Sentinel imagery using deep learning and large, globally distributed training sets [2, 9, 30]. Dynamic World predicts nine land-cover classes in near real time from Sentinel-2, with labels obtained by visual interpretation of time series and high-resolution basemaps, while Esri and WorldCover generate annual maps using supervised models trained on human-labelled Sentinel-1/2 samples. These products are extremely valuable as globally consistent baselines, but their legends use broad “crops”, “grass/shrub” or “rangeland” classes and do not explicitly separate scrublands, rain-fed smallholder farms, and agro-horticultural plantations. Comparative validation studies show that all three products perform relatively poorly globally for scrubland class. [18, 28].

**Grassland and pasture mapping.** The WRI Global Pasture Watch map focuses specifically on grasslands and pastures, using Landsat-based features and spatio-temporal machine learning to map cultivated and natural/semi-natural grasslands at 30 m resolution from globally distributed reference samples which was manually annotated by referencing very high resolution (VHR) imagery [21]. This improves the delineation of grass- and pasture-dominated ecosystems relative to generic global LULC products, but still does not resolve the finer land-use distinctions that are crucial in India, such as scrublands versus low-input croplands or plantations, and operates at coarser spatial resolution than the 10 m products. Across Dynamic World, Esri, WorldCover, and Pasture Watch, VHR imagery is primarily used for manual or point-based labelling.

**High-resolution segmentation and boundary detection.** A complementary body of work develops segmentation models for parcel or object delineation on high-resolution imagery. U-Net style encoder-decoder architectures and their variants (e.g., FracTAL-ResUNet) have been used to delineate smallholder field boundaries in India and other regions by predicting field interiors and boundary probability maps from 1.19 m satellite imagery [24] [26]. In parallel,

233 instance-segmentation extensions of YOLO[10] have been applied  
 234 to detect and segment objects such as buildings, tree crowns, or  
 235 plantation blocks in aerial and satellite images [31], and recent  
 236 foundation models like Segment Anything [23](SAM) have shown  
 237 promise for generic segmentation on remote-sensing data after  
 238 light adaptation or prompting [11] [20]. These approaches demon-  
 239 strate that VHR imagery can be used to recover clean, object-level  
 240 geometries for fields, scrub patches, or plantations, but most appli-  
 241 cations remain confined to limited study regions and are treated  
 242 as end products rather than as a mechanism for generating large,  
 243 geographically diverse training datasets for medium-resolution  
 244 classifiers.

245 **Positioning of this work.** Our contribution lies at the inter-  
 246 section of these strands. Unlike global 10 m products that learn  
 247 directly from medium-resolution samples and broad class legends,  
 248 we first use an information-theoretic sampling scheme to select  
 249 representative tiles across India’s agro-ecological zones, then apply  
 250 high-resolution segmentation and detection (FracTAL-ResUNet-  
 251 style boundary delineation and a YOLO-based plantation model) on  
 252 1.19 m imagery to automatically extract parcel-level scrub, farm, and  
 253 plantation geometries. These labels are then to train AEZ-specific  
 254 classifiers at 10 m resolution imagery. To our knowledge, this is one  
 255 of the first national-scale LULC pipelines that systematically uses  
 256 HR imagery as the primary source of algorithmically generated  
 257 training data.

### 260 3 Methodology

262 Distinguishing scrublands from rain-fed farmlands at national scale  
 263 requires far more than a single classifier. It demands a way to  
 264 systematically acquire high-quality training data across India’s  
 265 extraordinarily diverse agro-ecological regions. Our methodology  
 266 is therefore designed as a modular, scalable pipeline that begins by  
 267 identifying where representative information should come from,  
 268 then extracts reliable high-resolution labels, and finally trains a set  
 269 of region-specific mid-resolution classifiers capable of generalizing  
 270 across the country.

271 At a high level, the framework comprises five sequential compo-  
 272 nents Fig. 1. Starting with India’s 20 agro-ecological zones (AEZs),  
 273 the pipeline first searches for spatially and spectrally representa-  
 274 tive regions within each zone. High-resolution imagery from these  
 275 regions is then downloaded and processed using computer-vision  
 276 techniques to delineate farm, scrubland, and plantation boundaries.  
 277 These boundaries are then subjected to an information-theoretic  
 278 refinement stage that evaluates each segment using geometric,  
 279 contextual, and entropy-based criteria. By analysing attributes such as  
 280 shape regularity, spectral homogeneity, size thresholds, and bound-  
 281 ary entropy, the pipeline retains only those segments that exhibit  
 282 high internal consistency and clear land-cover identity to ensures  
 283 that the resulting training samples are high-confidence and spatially  
 284 reliable. Finally, a set of localized classifiers are trained on these di-  
 285 verse, AEZ-specific samples to produce pan-India predictions using  
 286 10 m satellite embeddings. The outputs from all AEZs are merged  
 287 to create a unified, national-scale LULC layer with significantly  
 288 improved class separation.

289 The five components of the pipeline are summarized below:

290 2026-02-09 11:36. Page 3 of 1–16.

- (1) **Spatial Representativeness and Tile Selection** – For  
 291 each AEZ, we identify tiles that best capture the zone’s  
 292 spectral, ecological, and geographic variability. This ensures  
 293 that subsequent sample generation draws from regions that  
 294 are truly representative of India’s diverse landscapes.  
 295
- (2) **High-Resolution Boundary Delineation** – Using 1.19  
 296 m imagery, we apply computer-vision segmentation to de-  
 297 delineate coherent land-parcel boundaries corresponding to  
 298 farms, scrublands, and plantation patches.  
 299
- (3) **Entropy and Geometry based Boundary Refinement**  
 300 – To achieve high-precision labels, we apply geometric,  
 301 morphological, and contextual rules that filter out noisy,  
 302 fragmented, or ambiguous boundaries, retaining only seg-  
 303 ments that can serve as trustworthy training samples.  
 304
- (4) **Sample Generation and Classifier Training** – Refined  
 305 segments are converted into labeled points that represent  
 306 the underlying class at 10 m resolution. These AEZ-specific  
 307 samples are used to train localized classifiers that capture  
 308 region-dependent spectral and seasonal behavior.  
 309
- (5) **Integration into the IndiaSAT Framework** – Classifier  
 310 outputs from all AEZs are harmonized within the IndiaSAT  
 311 pipeline to generate a cohesive pan-India scrubland, farm-  
 312 land, plantation layer, seamlessly integrated with existing  
 313 land-cover classes.  
 314

315 Together, these components form a scalable approach for pro-  
 316 ducing reliable training datasets across India’s heterogeneous land-  
 317 scapes—enabling mid-resolution models to learn distinctions that  
 318 would otherwise require infeasible amounts of manual labeling. De-  
 319 tailed descriptions of each component are provided in the following  
 320 subsections.  
 321

### 322 3.1 Spatial Representativeness and Tile 323 Selection

325 A central design choice in our pipeline is the use of high-resolution  
 326 (1.19 m) imagery which provides the spatial detail required for  
 327 computer-vision algorithms to extract clean, object-level segments  
 328 for the target land-use classes. Smallholder farms, scrub patches,  
 329 and plantation blocks, which are common across India, are visu-  
 330 ally distinguishable only at this scale. However, processing high-  
 331 resolution imagery uniformly across India’s 3.2 million km<sup>2</sup> land-  
 332 scape is computationally infeasible. We therefore needed a prin-  
 333 cipled method to identify a minimal subset of locations that still  
 334 captures the full geographical and spectral diversity of each AEZ.

335 To achieve this, we represent each AEZ as a collection of grids,

$$G = \{g_1, g_2, \dots, g_N\}, \quad (1)$$

336 where each grid spans a 9 km × 9 km region (32×32 tiles at 256 × 256  
 337 pixels, zoom level 17). This grid size is intentionally chosen: it is  
 338 large enough to capture the heterogeneous vegetation mosaics of  
 339 farms, scrublands, and plantation parcels while still remaining com-  
 340 putationally tractable for downstream segmentation. Our objective  
 341 is to identify a small subset  $S \subset G$  of size  $p$ , where  $p \ll N$ , such  
 342 that the selected  $p$  grids collectively represent the full spectral-  
 343 structural diversity of the AEZ. In effect, we aim to compress the  
 344 AEZ into a minimal, information-rich set of locations that can serve  
 345 as proxies for the entire region.  
 346

347 348

349 350

351

352 353

354

355 356

357

358 359

360

361 362

363

364 365

366

367 368

369

370 371

372

373 374

375

376 377

378

379 380

381

382 383

384

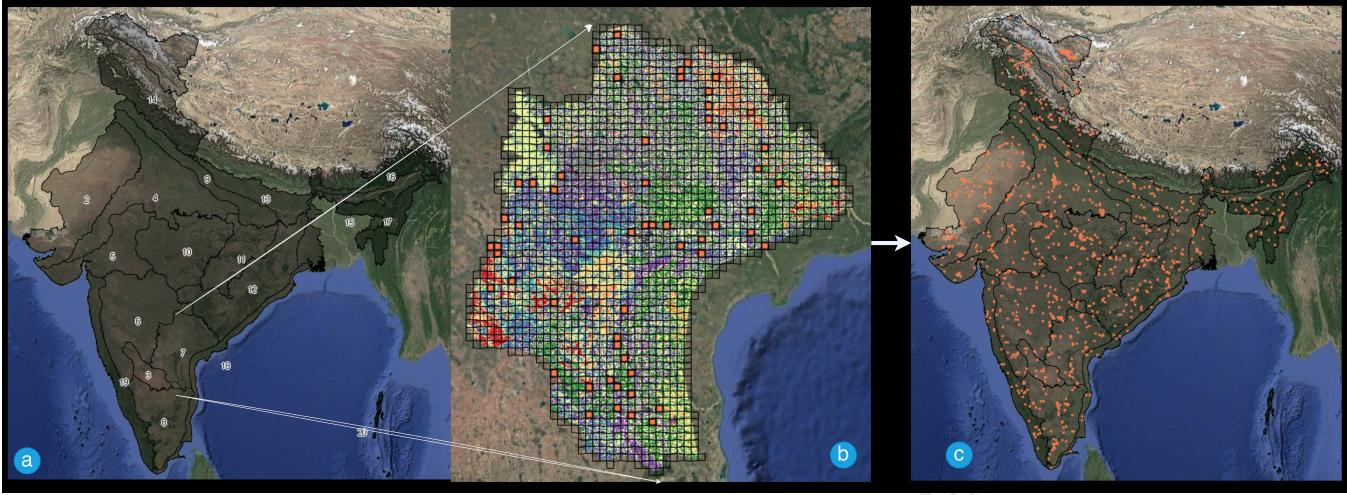


Figure 2: Figure showing representative tile selection across India. a) India divided into agro-ecological zones (AEZs) b) Here we show for AEZ 7, the grids, the clusters and the selected grids in orange c) Shows representative tiles selected from accross India

To meaningfully compare grids and decide which ones should be selected, each grid must first be represented by a compact descriptor that summarizes its vegetation and landscape characteristics. For this purpose, we construct a representative feature vector for every grid by leveraging Google’s 64-dimensional Alpha embeddings[4], which encode textural, structural, and spectral cues present in the underlying imagery.

We apply  $k$ -means clustering on all embeddings within the AEZ to partition the landscape into  $k$  characteristic vegetation–structure clusters. Each cluster roughly corresponds to a specific land-cover pattern (e.g., dense vegetation, sparse scrub, patchy cropland). For every grid  $g_i$ , we compute a normalized distribution  $P_i$  over these  $k$  clusters, capturing the proportion of each vegetation type found within that grid. This converts each grid into a comparable statistical signature.

To focus exclusively on relevant regions, we mask out urban, water, and barren areas using the IndiaSAT LULC layer before clustering so that cluster distributions reflect only farms, scrublands, and plantations.

Our goal is to select a subset

$$S \subset G, \quad |S| = p, \quad (2)$$

that is small enough to make high-resolution processing feasible yet diverse enough to represent the entire AEZ. A natural way to formalize this representativeness requirement is to compare the *aggregate* cluster distribution of the selected subset with the *overall* AEZ distribution. Specifically, if each grid  $g_i$  is associated with a cluster distribution  $P_i$ , then the subset-level distribution

$$P_S = \frac{1}{|S|} \sum_{g_i \in S} P_i \quad (3)$$

should closely approximate the AEZ-wide distribution

$$P_{AEZ} = \frac{1}{N} \sum_{g_i \in G} P_i. \quad (4)$$

By selecting the subset  $S$  whose aggregated distribution resembles  $P_{AEZ}$ , we ensure that the chosen grids collectively capture the same mixture of vegetation–structure patterns present across the full AEZ. In other words, the subset acts as a *compressed surrogate* for the entire zone, retaining its spectral and structural diversity while drastically reducing the volume of high-resolution imagery that must be processed downstream. This formulation provides a principled way to choose a compact but information-rich set of grids suitable for subsequent segmentation and sample generation.

Selecting the optimal subset  $S$  requires choosing the  $p$  grids whose combined distribution best matches the AEZ-wide distribution. However, evaluating all  $\binom{N}{p}$  subsets is computationally infeasible even for moderate  $N$ . To address this, we adopt a *greedy divergence-minimization strategy* that incrementally constructs a representative subset.

At each iteration, we evaluate every candidate grid  $g_i \in G \setminus S$  by measuring how much it improves the match between the subset distribution  $P_S$  and the overall AEZ distribution  $P_{AEZ}$ . Specifically, we use the Jensen–Shannon (JS) divergence[16]—a symmetric, information-theoretic measure of similarity between probability distributions—to quantify how representative the subset becomes when  $g_i$  is added. The algorithm selects the grid that yields the largest reduction in JS divergence:

$$g^* = \arg \min_{g_i \in G \setminus S} D_{JS}(P_{AEZ} \| P_{S \cup \{g_i\}}), \quad (5)$$

where the JS divergence is defined as

$$D_{JS}(P \| Q) = \frac{1}{2} D_{KL}(P \| M) + \frac{1}{2} D_{KL}(Q \| M), \quad M = \frac{1}{2}(P + Q), \quad (6)$$

and  $D_{KL}$  denotes the Kullback–Leibler divergence[13].

This iterative process continues until  $|S| = p$ , typically corresponding to only 3% of the grids within an AEZ as shown in Fig.. 2. Despite being small, this subset closely preserves the AEZ’s spectral–structural diversity, making it an efficient surrogate for

465 the entire zone. These selected grids then serve as the foundation  
 466 for high-resolution boundary delineation, ensuring that subsequent  
 467 computer-vision processing faithfully reflects India’s agro-  
 468 ecological heterogeneity while remaining computationally tractable.  
 469

470 Once the top- $p$  grids are selected, each  $32 \times 32$  grid is further  
 471 subdivided into four  $16 \times 16$  grids, which act as the atomic processing  
 472 units in the on-premise segmentation pipeline. The algorithm is  
 473 demonstrated in Appendix.

### 475 3.2 High-Resolution Boundary Delineation

476 The representative grids selected in the previous stage cover India’s full agro-ecological diversity, but manually annotating every high-resolution tile they contain would be prohibitively slow,  
 477 labor-intensive, and economically infeasible. Rather than relying  
 478 on human annotators to label millions of such parcels, we leverage  
 479 computer vision to extract these boundaries automatically, allowing  
 480 ground-truth creation to scale to the size and heterogeneity of the  
 481 Indian subcontinent.

482 For each selected grid subdivision, we download high-resolution  
 483 RGB imagery (1.19 m, zoom level 17)—a resolution at which India’s  
 484 smallholder farmland patterns become visually identifiable. Over  
 485 these tiles, we apply computer-vision models that segment the  
 486 imagery into coherent land-parcel objects corresponding to farms,  
 487 scrublands, and plantations. Once segments are extracted, we compute  
 488 a suite of geometric and textural descriptors for each object,  
 489 including area, entropy, rectangularity, and other properties that  
 490 capture the internal structure and visual coherence of the segment.  
 491 We download tiles at 4096X4096 pixels and then chunk it to images  
 492 of size 256X256 which is explained in appendix in more detail.  
 493

494 **3.2.1 Farm and Non-Agricultural Boundary Generation.** A key step  
 495 in our pipeline is the segmentation of high-resolution imagery into  
 496 meaningful land-parcel objects. We leverage an existing boundary-  
 497 delineation model which was created on a dataset of more than  
 498 10,000 annotated field boundaries in India and trained a FracTAL-  
 499 [26] specifically for smallholder agricultural landscapes. Their work  
 500 directly targets the visual characteristics of Indian farms, making  
 501 their model an ideal starting point and eliminating the need to  
 502 replicate expensive annotation and model-training efforts.

503 We apply this pretrained FracTAL-ResUNet to the  $256 \times 256$   
 504 patches. On agricultural regions, the model performs remarkably  
 505 well: it cleanly delineates the geometry of smallholder fields and pro-  
 506 duces precise boundary probability maps. However, when applied  
 507 to mixed landscapes that include scrublands, fallows, or degraded  
 508 non-agricultural areas, the model produces approximate “blobs.”  
 509 While these are not true field boundaries, they exhibit meaningful  
 510 structure—their shape irregularity, size, entropy, and rectangularity  
 511 differ systematically from those of real farm parcels.

512 This observation motivates a simple but powerful idea: although  
 513 the model was trained only on agricultural areas, its output still  
 514 contains enough spatial signal to separate field-like segments from  
 515 scrub or non-agricultural regions. By computing local geometric  
 516 and textural descriptors (e.g., area, perimeter, spectral entropy, rect-  
 517 angularity) for every segment produced by the model, we can sort  
 518 segments into three buckets: *farms*, *scrublands/non-agricultural*,

519 and *uncertain/other*. Even a small number of high-confidence seg-  
 520 ments in the first two categories is sufficient to generate reliable  
 521 ground-truth samples for training our downstream 10 m classifier.  
 522

523 *FracTAL-ResUNet Overview.* The segmentation backbone follows  
 524 the FracTAL-ResUNet architecture introduced by Waldner *et al.*,  
 525 which extends a U-Net encoder-decoder with skip connections by  
 526 replacing standard residual blocks with FracTAL attention units.  
 527 These units enhance the learning of fine boundary structures, mak-  
 528 ing the architecture highly effective for delineating smallholder  
 529 fields. The network performs multitask prediction, generating (i)  
 530 a field-extent mask, (ii) a boundary-probability map, and (iii) a  
 531 distance-to-boundary surface. This joint formulation significantly  
 532 improves boundary sharpness and parcel completeness.

533 Since the model outputs continuous boundary probability maps  
 534 rather than discrete polygons, we convert these to vector bound-  
 535 aries using a hierarchical watershed segmentation. The watershed  
 536 algorithm interprets the boundary-probability map as a topographic  
 537 surface and “floods” it from local minima, with ridges forming parcel  
 538 borders. We use the `higra` implementation, following the procedure  
 539 described in the original work.

540 A practical consideration is tile size. Although FracTAL-ResUNet  
 541 operates on  $256 \times 256$  inputs, watershed segmentation benefits from  
 542 running on a larger spatial context. For this reason, we first assem-  
 543 ble the  $16 \times 16$  grid of model outputs (boundary probability and  
 544 extent predictions) into a single large raster and run the watershed  
 545 algorithm at the block level. This design choice reflects both accu-  
 546 racy considerations and on-premise compute limitations: a  $16 \times 16$   
 547 block is the largest area we can process while ensuring tractable  
 548 memory usage.

549 We compute a set of geometric and textural descriptors for each  
 550 segment and store at this stage only. A detailed description of how  
 551 these features are computed and applied follows in the next section.

552 **3.2.2 Plantation Boundary Generation.** To demonstrate that our  
 553 pipeline can be extended beyond farms and scrublands to incorpo-  
 554 rate novel land-use classes with minimal additional effort, we de-  
 555 velop a dedicated module for detecting plantations. Unlike agricultural  
 556 fields, for which a high-quality pretrained boundary-delineation  
 557 model already exists, no analogous model was available for planta-  
 558 tion structures. Plantation blocks in India exhibit distinctive spatial  
 559 regularity and canopy patterns, but they vary widely by region and  
 560 are sparsely distributed, making manual annotation prohibitively  
 561 time-consuming. We therefore trained a dedicated deep-learning  
 562 model that operates on the same  $256 \times 256$  high-resolution tiles used  
 563 for farm segmentation, ensuring consistency and seamless integra-  
 564 tion into our pipeline. The entire training process is demonstrated  
 565 in appendix.

### 573 3.3 Entropy and Geometry Based Boundary 574 Refinement

575 The purpose of this refinement module is twofold: (i) to sepa-  
 576 rate the mixed FracTAL-ResUNet segments into high-confidence  
 577 farms, high-confidence Non-agro land, and uncertain regions using  
 578 entropy- and geometry-based criteria, and (ii) to prune noisy or  
 579

581 low-confidence plantation detections. The resulting refined bound-  
 582 aries form the set of reliable parcel geometries from which training  
 583 samples can be drawn.

584 *3.3.1 Farm and Non-Agricultural Boundary Refinement.* Our key  
 585 insight is that farmland parcels—being human-designed and sys-  
 586 tematically cultivated—tend to exhibit much lower randomness  
 587 and far greater structural regularity than scrublands or other non-  
 588 agricultural formations. Scrub patches, by contrast, are shaped by  
 589 ecological processes and therefore display irregular geometry and  
 590 higher textural complexity. We quantify these differences using a  
 591 set of descriptive properties such as entropy, rectangularity, and  
 592 size. These measures, and the intuition behind why they separate  
 593 farms from non-agricultural parcels, are described below before  
 594 being formalized into the refinement rules used by our pipeline.  
 595

596 **Entropy:** Local textural complexity provides a strong cue for  
 597 distinguishing human-made farmland parcels from naturally oc-  
 598 curring scrublands. Scrub regions are shaped by ecological pro-  
 599 cesses—patchy shrubs, scattered bushes, irregular canopies—and  
 600 therefore exhibit highly heterogeneous textures. In contrast, agri-  
 601 cultural fields are typically more uniform: ploughed soil, single-crop  
 602 canopies, or bare field interiors, interrupted only occasionally by an  
 603 isolated tree or small vegetation patch. This fundamental difference  
 604 makes entropy a powerful discriminator.

605 We quantify local spatial complexity using per-pixel Shannon en-  
 606 tropy computed with the entropy function from the `scikit-image`  
 607 library [12]. Entropy is evaluated within a disk-shaped neighbor-  
 608 hood of radius  $r$ , where the normalized histogram of gray-level  
 609 intensities,  $p_i$ , defines:

$$610 \quad H = - \sum_{i=1}^L p_i \log_2(p_i), \quad (7)$$

611 with  $L$  denoting the number of intensity levels. Homogeneous zones  
 612 produce entropy values close to zero, whereas scrub patches—with  
 613 their intricate, irregular textures—produce consistently high en-  
 614 tropy responses. We use  $r = 5$ , which empirically balances noise  
 615 sensitivity and spatial stability.

616 A practical nuance is that boundary pixels around farm parcels  
 617 often have artificially high entropy due to sharp edges, crop borders,  
 618 shadows, and segmentation imperfections. If used directly, these  
 619 boundary effects would inflate entropy for otherwise clean agricul-  
 620 tural fields. To avoid this, we focus on the interior of each segment.  
 621 Let  $\mathcal{S}$  be all pixels in a segment, and define the high-entropy core  
 622 as:

$$623 \quad S^+ = \{x \in \mathcal{S} \mid H(x) > 5.2\},$$

624 where the threshold isolates genuinely heterogeneous regions while  
 625 excluding noisy boundary artifacts.

626 The segment-level entropy score is then:

$$627 \quad \bar{H}_{\mathcal{S}} = \frac{1}{|\mathcal{S}^+|} \sum_{x \in \mathcal{S}^+} H(x). \quad (8)$$

628 By emphasizing only the mid-region of each segment,  $\bar{H}_{\mathcal{S}}$  be-  
 629 comes a robust indicator of natural heterogeneity: scrubland seg-  
 630 ments retain consistently high values, whereas farmland segments—aside  
 631 from small local variations such as a single tree—exhibit markedly  
 632 lower entropy. This contrast enables reliable separation between  
 633 farm and non-agricultural regions.

634 **Rectangularity Measure:** Shape regularity provides another  
 635 strong cue for separating farmland parcels from naturally formed  
 636 scrub regions. Agricultural fields in India are typically shaped by hu-  
 637 man planning, past land consolidation, or mechanized cultivation,  
 638 and therefore tend to approximate simple geometric forms—most  
 639 often rectangles or rectilinear polygons. Scrublands, by contrast,  
 640 arise through ecological processes and exhibit meandering, branch-  
 641 ing, and highly irregular outlines. Rectangularity captures this  
 642 distinction directly.

643 Let  $\mathcal{S}$  denote the set of pixels belonging to a segment. We first  
 644 extract its external contour using OpenCV’s `cv2.findContours`,  
 645 yielding an area  $A_{\text{contour}}$  corresponding to the true region footprint.  
 646 We then compute the minimum-area rotated bounding rectangle  
 647 that encloses this contour using `cv2.minAreaRect`, and denote its  
 648 area by  $A_{\text{rect}}$ . The rectangularity score is defined as:

$$649 \quad R = \frac{A_{\text{contour}}}{A_{\text{rect}}}, \quad (9)$$

650 with values constrained to the interval  $[0, 1]$ . A perfect rectangle  
 651 achieves  $R = 1$ , while deviations from rectilinearity—indentations,  
 652 concave boundaries, or curved edges—reduce the ratio.

653 Rectangularity thus serves as a direct indicator of human inter-  
 654 vention: high values correspond to well-shaped farmland parcels,  
 655 whereas low values reflect the organic, irregular geometries char-  
 656 acteristic of scrublands and other non-agricultural regions. When  
 657 combined with entropy, it complements textural information by pro-  
 658 viding a robust shape-based criterion for distinguishing cultivated  
 659 land from natural vegetation mosaics.

660 **Size:** The physical extent of a segmented region also provides  
 661 a useful cue for distinguishing agricultural fields from scrublands.  
 662 In our study areas, typical smallholder farm parcels fall within the  
 663 range of approximately  $500\text{--}2000 \text{ m}^2$ , reflecting dominant regional  
 664 agricultural patterns. By contrast, scrubland segments are generally  
 665 much larger and more spatially diffuse. Segment size therefore acts  
 666 as a coarse but effective discriminator when combined with textural  
 667 (entropy) and geometric (rectangularity) properties.

668 *Removing Isolated Farm-Like Artifacts.* When applying these  
 669 rules directly, we observed that some isolated segments located  
 670 inside non-agricultural regions also satisfy the entropy, rectangu-  
 671 larity, and size thresholds. Such isolated detections act as noise when  
 672 used for training and therefore need to be removed. We found  
 673 that genuine farms tend to occur in spatial clusters, whereas these  
 674 spurious detections almost always appear as single, isolated shapes.

675 To address this, we constructed spatial clusters of farm candi-  
 676 dates by grouping segments whose nearest boundaries lie within  
 677  $10 \text{ m}$  of each other. For each cluster, we counted the number of farm  
 678 candidates it contained. Only clusters with three or more segments  
 679 were retained, and all farm candidates within clusters of size  $\geq 3$   
 680 were accepted as valid. Isolated segments (clusters of size 1 or 2)  
 681 were removed.

682 *Final Rules:* A high-confidence farm boundary will exhibit fol-  
 683 lowing rules:

- (1) Entropy  $< 1.0$ ,
- (2) Rectangularity  $> 0.67$ ,
- (3) Size  $\in [500, 2000] \text{ m}^2$ ,

- 697 (4) Membership in a spatial cluster containing at least three  
 698 farm candidates.

699 These criteria yield a small but highly reliable set of farm bound-  
 700 aries, suitable for generating clean training samples for the down-  
 701 stream classifier. Check appendix for further explanation.

702 **3.3.2 Non-Agricultural Area Identification Rules.** To identify high-  
 703 confidence scrubland and non-agricultural segments, we found that  
 704 segment size provides the most reliable and consistent discriminator  
 705 across all Agro-Ecological Zones (AEZs). In every region examined,  
 706 boundaries corresponding to scrublands were substantially larger  
 707 than those of agricultural fields. Empirically, scrubland segments  
 708 almost always exceeded 60,000 m<sup>2</sup>. We therefore use segment size  
 709 as the primary criterion for selecting non-agricultural boundaries.  
 710 To avoid impractically large polygons—which may cause memory  
 711 overflow during later validation steps—we cap the upper size at  
 712 5,000,000 m<sup>2</sup>.

713 A second filtering step is required to prevent the accidental inclu-  
 714 sion of agricultural regions. We observed that the FracTAL-ResUNet  
 715 model occasionally produces very large segments over burnt fields  
 716 or ploughed-off-season croplands. Because these boundaries are  
 717 both large and visually irregular, they satisfy the scrubland size cri-  
 718 terion and are mistakenly placed into the non-agricultural bucket.  
 719 To remove such false positives, we cross-validate each candidate  
 720 segment against the IndiaSAT v3 which is our inhouse LULC layer.  
 721 Segments in which more than 50% of pixels correspond to agricul-  
 722 ture are excluded, effectively filtering out burnt-field artifacts and  
 723 other misclassified agricultural areas.

724 A non-agricultural boundary exhibits following rules:

- 725 (1) Size  $\in [60,000, 5,000,000]$  m<sup>2</sup>,  
 726 (2) IndiaSAT v3 cross-validation: > 50% non-agricultural pixels,  
 727 removing burnt-field artifacts generated by the FracTAL-  
 728 ResUNet model.

729 **3.3.3 Plantation Filtering Rules.** We observed that, without addi-  
 730 tional filtering, the Yolo model occasionally identified overly large  
 731 forested regions as plantations or fragmented tree patches as plan-  
 732 tation units. To ensure that only realistic plantation blocks were  
 733 retained, we applied a size-based filter. Only those segments were  
 734 accepted which has their area within the range [1,000, 20,000] m<sup>2</sup>.  
 735 The lower bound removes scattered tree patches or small isolated  
 736 detections that do not correspond to structured plantation rows,  
 737 while the upper bound filters out large, continuous forest regions  
 738 misinterpreted as single plantation blocks.

739 A plantation boundary exhibits following rules:

- 740 (1) Area  $\in [1,000, 20,000]$  m<sup>2</sup>, removing scattered trees and  
 741 oversized forest-like detections.

### 742 **3.4 Sample Generation and Classifier Training**

743 Once high-confidence boundaries were identified, we generated  
 744 training samples for each Agro-Ecological Zone (AEZ). For every  
 745 validated segment, uniformly distributed sample points were ex-  
 746 tracted, ensuring that each 16 × 16 grid produced approximately  
 747 150 samples per class (farm, non-agricultural land, and plantation),  
 748 wherever those classes were present. Overall, sampling covered  
 749 nearly 3% of the total AEZ area, providing a spatially diverse yet  
 750 computationally manageable dataset.

752 To classify the entire AEZ at 10 m resolution, we relied on Google’s  
 753 64-dimensional embedding vectors, which encode annual spectral-  
 754 temporal information from multiple satellite sensors. For every  
 755 sampled point, we extracted embedding vectors for the past three  
 756 years. This assumes that land use remains stable over this period  
 757 for our three target classes: farms, non-agricultural land, and plan-  
 758 tations. Multi-year embeddings also offer an important practical  
 759 advantage: they capture temporal context for younger plantations.  
 760 Areas that are mature plantations today may have appeared as  
 761 young stands in earlier years.

762 A separate Random Forest classifier was then trained for each  
 763 AEZ using the multi-year embeddings. This AEZ-specific approach  
 764 allows the model to adapt to regional ecological and agricultural  
 765 conditions, avoiding the over-generalization that occurs when train-  
 766 ing a single nationwide model. The resulting pool of AEZ-level  
 767 classifiers was subsequently applied across each zone to generate  
 768 full-coverage maps at 10 m spatial resolution.

769 Finally, predictions from all AEZs were combined to produce  
 770 a pan-India land-use map containing three classes—farm, non-  
 771 agricultural land, and plantation—derived from high-confidence  
 772 samples, regionally tuned models, and multi-year embedding fea-  
 773 tures.

### 774 **3.5 Integration into the existing LULC 775 Framework**

776 To create a complete, hierarchical LULC map, the output of our  
 777 classifier was integrated into the IndiaSAT v3 pipeline as follows:

- 778 (1) Initialized background (0) and added built-up (1), water (2),  
 779 and barren (3) pixels using IndiaSAT v3 modules.  
 780 (2) For remaining pixels, applied our classifier output to divide  
 781 regions into farms, non-agro, and plantations (13).  
 782 (3) Further refined:  
 783   • Farm regions split into four crop-intensity classes—single  
 784     Kharif (8), single non-Kharif (9), double (10), triple (11).  
 785   • Non-agro regions subdivided using the IndiaSAT tree  
 786 classifier into forest and non-forest areas; the non-  
 787 forest areas were designated as scrublands (12).

788 This multi-stage integration produced a refined LULC layer.

## 789 **4 Results**

790 We evaluated the performance of our LULC framework using mul-  
 791 tiple publicly available datasets. A key limitation in benchmarking  
 792 global LULC models is the absence of a universally accepted, inde-  
 793 pendently held-out validation dataset. Most global LULC validation  
 794 either (i) use proprietary reference datasets that are not publicly  
 795 released, or (ii) rely on subsets of the Dynamic World (DW) Test  
 796 dataset for evaluation, which restricts full cross-product compara-  
 797 bility. WC used a completely different dataset to train and test their  
 798 model and so this dataset can be used for validating ours with DW  
 799 and WC. ESRI Global LULC model explicitly uses the DW dataset  
 800 for training, without clarifying whether the DW Test subset was  
 801 excluded during training. This ambiguity prevents DW Test from  
 802 serving as a fully independent benchmark for ESRI. Nevertheless,  
 803 because DW Test is the only publicly accessible dataset containing  
 804 both scrub and cropland classes for India, we evaluate our frame-  
 805 work on it. To further strengthen the validity of our evaluation  
 806 807 808 809 810 811 812

Dataset	Ours	WorldCover	Dynamic World	ESRI Land Cover	WRI
DW Test data[3]	OA = <b>78.13%</b> PA <sub>S</sub> = 20.08%, UA <sub>S</sub> = 20.37% PA <sub>F</sub> = <b>81.28%</b> , UA <sub>F</sub> = 98.60%	OA = 64.50% PA <sub>S</sub> = 6.01%, UA <sub>S</sub> = 9.32% PA <sub>F</sub> = 67.67%, UA <sub>F</sub> = 99.07%	OA = 45.76% PA <sub>S</sub> = 38.34%, UA <sub>S</sub> = 16.12% PA <sub>F</sub> = 46.17%, UA <sub>F</sub> = 98.05%	OA = 75.52% PA <sub>S</sub> = <b>54.61%</b> , UA <sub>S</sub> = <b>26.95%</b> PA <sub>F</sub> = 76.69%, UA <sub>F</sub> = <b>99.39%</b>	–
WRI Dataset[21]	<b>95.67%</b>	82%	90%	84%	93.54%
10000 Fields[25]	<b>90.82%</b>	89.96%	79.53%	85.37%	–
AgriFieldNet[6]	<b>95.29%</b>	84.72%	72.70%	83.50%	–

Table 1: Evaluation of our LULC against global LULC products across multiple categories. PA denotes Producer Accuracy, UA denotes User Accuracy, S denotes Scrub, and F denotes Farms.

and explicitly use dataset not employed for training we used three independent references datasets which are not meant for validation but could be used. We explain our datasets below:

- (1) DW Test dataset is a manually annotated dataset and is a set of test tiles held aside from training. There are 409 validation tiles out of which 40 lies inside India. This dataset is already used in verifying global LULC here[paper].
- (2) WRI Global Pasture Watch dataset is in which they annotated grasslands both cultivated and Natural/semi-natural globally using VHR imager. We used the natural/semi natural grasslands class in this to test our Non-agro classification as most of their class occurs in our Non-agro class rather than completely lying in scrubland class. To test farm class independently, we used 10000 Fields dataset and AgrifieldNet dataset.
- (3) 10000 Fields dataset is a the only pan India fields dataset where they marked 5 fields from each reference regions amounting of annotating a total of 10000 fields in total.
- (4) Agrifield net is also another fields datset which comprises of fields in four states of Uttar Pradesh, Rajasthan, Odisha and Bihar in northern India.

Table 1 summarizes these comparisons. On the DW Test dataset, our model achieves the highest overall accuracy (78.13%), with particularly strong performance for the farmland class (PA = 81.28%, UA = 98.60%). The overall accuracy show that Our framework outperform in clearly seperating scrubland from farmlands. Scrubland remains challenging across all LULC products due to its structural and phenological similarity to plantations and mixed vegetation mosaics; ESRI achieves the highest scrub PA on DW (54.61%), likely reflecting overlap between its training data and the DW dataset. Performance on the WRI Pasture Watch dataset highlights the model’s ability to distinguish natural vegetation from agricultural areas. Our approach achieves 95.67% accuracy, outperforming all global products and demonstrating robust classification of non-agricultural landscapes, despite their ecological and spectral variability. Evaluation on field-level datasets further confirms the model’s superiority in agricultural regions. On the 10000 Fields dataset—the only pan-India field-boundary dataset—our method obtains the highest accuracy (90.82%), marginally exceeding WorldCover and substantially surpassing Dynamic World and ESRI. The trend is even more pronounced for AgriFieldNet, where our model achieves 95.29% accuracy. These results collectively demonstrate that India-specific ecological modeling and AEZ-aware generalization substantially outperform global LULC products in the Indian context.

Since the primary challenge addressed in this work is the absence of a high-quality, well-distributed, large-scale reference dataset for

India, the available public datasets cannot fully capture the extent of improvement achieved by our framework. To complement the quantitative evaluation, we additionally provide qualitative visual comparisons, showcasing representative examples where our method offers clearer boundaries, better class separation, and more consistent spatial patterns. Refer to the appendix for comparative diagrams illustrating our LULC outputs alongside global LULC products across multiple regions, clearly highlighting the improved class separation and reduced misclassification achieved by our approach.

Due to the lack of public plantation boundary datasets, we also conducted a qualitative validation using temporal imagery on **Google Earth Pro**[27]. We compared the plantation coverage generated by our model from 2017 to 2025 with historical imagery, visually assessing expansion or shrinkage trends. Our LULC outputs demonstrated consistent detection of plantation patches corresponding to visible canopy expansion in the high-resolution imagery. This validation confirms that the model effectively captures gradual plantation growth while maintaining minimal false detections in non-plantation regions. Refer to appendix for qualitative examples.

## 5 Discussion and Future Work

The CV-derived labels from high-resolution imagery remain an approximation of ground truth. Our entropy- and geometry-based rules deliberately trade off coverage for reliability: they discard many valid but complex parcels in order to retain a smaller set of high-confidence examples. A natural next step is to construct a compact, hand-annotated reference dataset focused explicitly on scrubland–farmland transitions and mixed mosaics. Because our representative-tile selection already identifies the most informative grids within each AEZ, these tiles provide a principled starting pool for manual annotation. Training models directly on such human-labelled data could further improve class separability and reduce residual biases in difficult landscapes.

Finally, the pipeline we propose is not restricted to the three classes studied here. The same strategy of tile selection, high-resolution segmentation, and AEZ-specific learning on 10 m embeddings—can be used to introduce additional, contextually important classes with relatively few examples, such as separating natural forests from tree plantations, or subdividing open natural ecosystems into grasslands and shrublands. We also expect that future work incorporating explicit uncertainty estimates, temporal consistency constraints, and cross-country transfer to other semi-arid regions could turn this approach into a general recipe for building regionally contextualised LULC products.

929

## References

- [1] M Arasumani, Danish Khan, Arundhati Das, Ian Lockwood, Robert Stewart, Ravi A Kiran, M Muthukumar, Milind Bunyan, and VV Robin. 2018. Not seeing the grass for the trees: Timber plantations and agriculture shrink tropical montane grassland by two-thirds over four decades in the Palani Hills, a Western Ghats Sky Island. *PLoS one* 13, 1 (2018), e0190003.
- [2] Christopher F Brown, Steven P Brumby, Brookie Guzder-Williams, Tanya Birch, Samantha Brooks Hyde, Joseph Mazzariello, Wanda Czerwinski, Valerie J Pasquarella, Robert Haertel, Simon Ilyushchenko, et al. 2022. Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific data* 9, 1 (2022), 251.
- [3] Christopher F. Brown, Steven P. Brumby, Brookie Guzder-Williams, Tanya Birch, Samantha Brooks Hyde, Joseph Mazzariello, Wanda Czerwinski, Valerie J. Pasquarella, Robert Haertel, Simon Ilyushchenko, Kurt Schwehr, Mikaela Weisse, Fred Stolle, Craig Hanson, Oliver Guinan, Rebecca Moore, and Alexander M. Tait. 2021. Dynamic World Test Tiles (Version v2). Zenodo. doi:10.5281/zenodo.4766508
- [4] Christopher F Brown, Michal R Kazmierski, Valerie J Pasquarella, William J Rucklidge, Masha Samsikova, Chenhui Zhang, Evan Shelhamer, Estefania Lahera, Olivia Wiles, Simon Ilyushchenko, et al. 2025. Alphaearth foundations: An embedding field model for accurate and efficient global mapping from sparse label data. *arXiv preprint arXiv:2507.22291* (2025).
- [5] Matthew E Fagan, Do-Hyung Kim, Wesley Settle, Lexie Ferry, Justin Drew, Haven Carlson, Joshua Slaughter, Joshua Schaferbien, Alexandra Tyukavina, Nancy L Harris, et al. 2022. The expansion of tree plantations across tropical biomes. *Nature Sustainability* 5, 8 (2022), 681–688.
- [6] Radian Earth Foundation and IDinsight. 2022. AgriFieldNet Competition Dataset (Version 1.0). Radiant MLHub. doi:10.34911/rdnt.wu92p1
- [7] Madhav Gadgil and Ramachandra Guha. 2013. *Ecology and equity: The use and abuse of nature in contemporary India*. Routledge.
- [8] Narpat S Jodha. 2000. Waste lands management in India: Myths, motives and mechanisms. *Economic and Political Weekly* (2000), 466–473.
- [9] Krishna Karra, Caitlin Kontgis, Zoe Statman-Weil, Joseph C Mazzariello, Mark Mathis, and Steven P Brumby. 2021. Global land use/land cover with Sentinel 2 and deep learning. In *2021 IEEE international geoscience and remote sensing symposium IGARSS*. IEEE, 4704–4707.
- [10] Rahima Khanam and Muhammad Hussain. 2024. Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725* (2024).
- [11] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. 2023. Segment anything. In *Proceedings of the IEEE/CVF international conference on computer vision*, 4015–4026.
- [12] Oliver Kramer. 2016. Scikit-learn. In *Machine learning for evolution strategies*. Springer, 45–53.
- [13] Solomon Kullback and Richard A Leibler. 1951. On information and sufficiency. *The annals of mathematical statistics* 22, 1 (1951), 79–86.
- [14] Sutirtha Lahiri and Sushma Reddy. 2025. India's grasslands are not "wastelands". *Science* 387, 6735 (2025), 726–727.
- [15] Sutirtha Lahiri, Anirban Roy, and Forrest Fleischman. 2023. Grassland conservation and restoration in India: A governance crisis. *Restoration Ecology* 31, 4 (2023), e13858.
- [16] Jianhua Lin. 2002. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information theory* 37, 1 (2002), 145–151.
- [17] AS MacDougal, B Vanzant, J Sulik, S Bagchi, D Naidu, TO Muraina, EW Seabloom, ET Borer, P Wilfahrt, I Slette, et al. 2026. The global extent of the grassland biome and implications for the terrestrial carbon sink. *Nature Ecology & Evolution* (2026), 1–12.
- [18] MD Madhusudan and Abi Tamim Vanak. 2023. Mapping the distribution and extent of India's semi-arid open natural ecosystems. *Journal of Biogeography* 50, 8 (2023), 1377–1387.
- [19] Daniel McGahey, Jonathan Davies, Nikolas Hagelberg, and Razingrim Ouedraogo. 2014. Pastoralism and the Green Economy—a natural nexus. *Nairobi: IUCN and UNEP. x+ 58p* (2014).
- [20] Lucas Prado Osco, Qiusheng Wu, Eduardo Lopes De Lemos, Wesley Nunes Gonçalves, Ana Paula Marques Ramos, Jonathan Li, and José Marcato Junior. 2023. The segment anything model (sam) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation* 124 (2023), 103540.
- [21] Leandro Parente, Lindsey Sloat, Vinicius Mesquita, Davide Consoli, Radost Stanimirova, Tomislav Hengl, Carmelo Bonannella, Nathália Teles, Ichsan Wheeler, Maria Hunter, et al. 2024. Annual 30-m maps of global grassland class and extent (2000–2022) based on spatiotemporal Machine Learning. *Scientific data* 11, 1 (2024), 1303.
- [22] RH Rizvi, AK Handa, KB Sridhar, A Kumar, S Bhaskar, SK Chaudhari, A Arunachalam, Noyal Thomas, S Ashutosh, RK Sapra, et al. 2020. Mapping agroforestry and trees outside forest. (2020).
- [23] Mélisande Teng, Arthur Ouaknine, Etienne Laliberté, Yoshua Bengio, David Rolnick, and Hugo Larochelle. 2025. Bringing SAM to new heights: Leveraging elevation data for tree crown segmentation from drone imagery. *arXiv preprint arXiv:2506.04970* (2025).
- [24] François Waldner, Foivos I Diakogiannis, Kathryn Batchelor, Michael Ciccotosto-Camp, Elizabeth Cooper-Williams, Chris Herrmann, Gonzalo Mata, and Andrew Toovey. 2021. Detect, consolidate, delineate: Scalable mapping of field boundaries using satellite images. *Remote sensing* 13, 11 (2021), 2197.
- [25] Sherrie Wang, Francois Waldner, and David B. Lobell. 2022. 10,000 Crop Field Boundaries across India. Zenodo. doi:10.5281/zenodo.7315090
- [26] Sherrie Wang, François Waldner, and David B Lobell. 2022. Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision. *Remote Sensing* 14, 22 (2022), 5738.
- [27] Dennis Wuthrich. 2006. Google Earth Pro. *Geospatial Solutions* 16, 2 (2006).
- [28] Panpan Xu, Nandin-Erdene Tsednbazar, Martin Herold, Sytze De Bruin, Myke Koopmans, Tanya Birch, Sarah Carter, Steffen Fritz, Myroslava Lesiv, Elise Mazur, et al. 2024. Comparative validation of recent 10 m-resolution global land cover maps. *Remote Sensing of Environment* 311 (2024), 114316.
- [29] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. 2019. Cutmix: Regularization strategy to train strong classifiers with localizable features. In *Proceedings of the IEEE/CVF international conference on computer vision*, 6023–6032.
- [30] Daniele Zanaga, Ruben Van De Kerchove, Dirk Daems, Wanda De Keersmaecker, Carsten Brockmann, Grit Kirches, Jan Wevers, Oliver Cartus, Maurizio Santoro, Steffen Fritz, et al. 2022. ESA WorldCover 10 m 2021 v200. (2022).
- [31] Juepeng Zheng, Weijia Li, Maocai Xia, Runmin Dong, Haohuan Fu, and Shuai Yuan. 2019. Large-scale oil palm tree detection from high-resolution remote sensing images using faster-rcnn. In *Igarss 2019-2019 ieee international geoscience and remote sensing symposium*. IEEE, 1422–1425.

## A Appendix

### A.1 Algorithm 1

**Algorithm 1:** Greedy Subset Selection for AEZ Representativeness

---

**Input:** Grids  $G = \{g_1, \dots, g_N\}$ , AEZ-level distribution  $P_{AEZ}$ , desired subset size  $p$

**Output:** Representative subset  $S$

```

 $S \leftarrow \emptyset$  // Initialize empty subset
while  $|S| < p$  do
    foreach  $g_i \in G \setminus S$  do
        Compute JS divergence after adding  $g_i$ :  

         $d_i = D_{JS}(P_{AEZ} \parallel P_{S \cup \{g_i\}})$ 
    Select the grid that minimizes divergence:  

     $g^* = \arg \min_{g_i \in G \setminus S} d_i$ 
    Update subset:  

     $S \leftarrow S \cup \{g^*\}$ 

```

**return**  $S$

---

### A.2 Plantation Training

To demonstrate that our pipeline can be extended beyond farms and scrublands to incorporate novel land-use classes with minimal additional effort, we develop a dedicated module for detecting plantations. Unlike agricultural fields, for which a high-quality pre-trained boundary-delineation model already exists, no analogous model was available for plantation structures. Plantation blocks in India exhibit distinctive spatial regularity and canopy patterns, but they vary widely by region and are sparsely distributed, making manual annotation prohibitively time-consuming. We therefore trained a dedicated deep-learning model that operates on the same  $256 \times 256$  high-resolution tiles used for farm segmentation, ensuring consistency and seamless integration into our pipeline.

987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044

1045 *Training Data and Model:* We curated a plantation dataset by collecting high-resolution imagery from four plantation-rich regions across Karnataka and Maharashtra: Tumakuru, Kolar, Kadiri, and Paithan. This process yielded 9,811 candidate image tiles. Using the Roboflow platform, we manually annotated a subset of these tiles, resulting in 657 training examples and 72 test examples containing a total of 1,215 individual polygon annotations (see Table 2).

1052 To address the inherent sparsity of plantations within the landscape, we fine-tuned a YOLOv11 model [10] pretrained on COCO and ImageNet datasets. The model was trained for 100 epochs with standard augmentations (random flips, rotations, translations) and optimized using a OneCycleLR scheduler. The resulting model achieved robust performance, with mAP<sub>50</sub> scores exceeding 0.83 for both bounding-box and mask predictions (see Figure ?? in Appendix).

1060 *Inference and Generalization:* The trained model was then applied to new plantation candidates across India using the same zoom level 17 imagery used during training. The model outputs polygonal masks for plantation blocks, with predictions filtered using a confidence threshold of 0.5 to ensure reliability. Predicted mask coordinates (in pixel space) were mapped to their geographic locations, enabling seamless integration into downstream geospatial processing. An example of plantations detected by our model is shown in Figure 3.

1070 *Reducing False Positives with CutMix Augmentation:* Initial inference revealed a tendency for the model to produce false positives in landscapes where natural tree clusters or non-plantation orchards mirrored the geometric regularity of managed plantations (see Figure 4). To mitigate this structural confusion, we implemented a targeted augmentation strategy based on the CutMix technique [29]. We curated 100 negative samples - high-resolution tiles devoid of plantations but containing problematic vegetation structures - from regions where false detections were most frequent. For each negative sample, we generated two unique synthetic training examples by copy-pasting distinct, authenticated plantation polygons into the negative context. This process expanded the training set from 657 to 857 images and increased the total annotation count to 1,415 (see Table 2). Retraining the YOLOv11 model with this spatially-hardened dataset significantly suppressed these false positives cases.

1086 *Integration into the Pipeline:* The final plantation model produces instance-level polygon masks at 1.19 m resolution, analogous to those obtained for farms and scrublands. These plantation boundaries are then merged with the outputs of the previous module and passed through the same statistical and geometric validation framework (Section ??). This demonstrates that our approach generalizes naturally to new land-use categories: once high-resolution segments are extracted, their boundaries can be validated, filtered, and converted into high-confidence samples for 10 m classification—without requiring any modification to the core pipeline. The complete workflow for data refinement, model training, and the inference architecture is illustrated in Figure 6.

### A.3 Downloading and Chunking

1100 To obtain high-resolution inputs for segmentation, we use the `tms_to_geotiff` utility from the `segment-geospatial` library, which



1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
**Figure 3: True Positives**



1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
**Figure 4: False Positives**

1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
**Figure 5: Examples of plantation detection results. Left (a): True positive cases showing correctly identified plantations. Right (b): False positive cases where the model incorrectly classified non-plantation structures as plantations.**

1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
**Table 2: Plantation Dataset details and CutMix impact.**

Dataset	Train	Test	Total	Annotations
Initial	657	72	729	1,215
Post-CutMix	857	72	929	1,415

1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
**Table 3: Mean Average Precision (mAP) results for object detection (Box) and image segmentation (Mask).**

Model	mAP50 (Box)	mAP50 (Mask)
Plantation Detection Model	0.83347	0.85314
Plantation Detection Model after CutMix Augmentation	0.83794	0.84968

1160  
downloads imagery from a Tile Map Service (TMS) and converts it into a GeoTIFF. For each selected grid subdivision, we retrieve a GeoTIFF at zoom level 17, corresponding to a  $16 \times 16$  tile block where each tile is a  $256 \times 256$  pixel image.

Since our computer-vision models operate on  $256 \times 256$  inputs, each downloaded GeoTIFF is further split (“chunked”) into individual  $256 \times 256$  patches. These patches serve as the atomic inputs for the segmentation pipeline, ensuring compatibility with the model architecture. We store the individual image position in their file address which we use in future to stitch them again into  $16 \times 16$  block of outputs.<sup>4</sup>



**Figure 6: Plantation Detection Workflow.** *Refinement and Training (Left):* Negative samples are hardened by "cutting and pasting" annotations from training data to repopulate the dataset. A YOLOv11 model is trained on a 90-10 split, utilizing a 5 × 5 grid on input to generate bounding boxes and class probability maps. *Inference (Right):* Zoom level 17 image tiles are processed through a convolutional backbone with multiple down-sampling layers. A projection head and dense network then compute the final detection, with predictions filtered by a confidence score > 0.5.

#### A.4 Overlap Resolution

Because farm and scrubland boundaries originate from the same FracTAL-ResUNet segmentation pipeline, they are mutually exclusive by construction. Plantation boundaries, however, are generated independently using the YOLO-based detection model and may spatially overlap with either farms or scrublands. Such overlaps would result in duplicate or conflicting labels if not resolved.

To maintain a clean and non-redundant set of training regions, we apply a simple hierarchical precedence rule:

##### Plantation > Farm > Scrubland

This hierarchy reflects the relative specificity and reliability of each boundary type. Whenever a plantation segment overlaps with a farm or scrubland boundary, the plantation label is retained and the underlying boundary is discarded. Similarly, in the rare cases where farm and scrubland segments touch or overlap, the farm label is preferred, consistent with the sharper geometric and textural cues associated with agriculture.

This precedence rule ensures that each spatial region contributes exactly one class label, preventing duplicate sampling and guaranteeing consistency across the training dataset.

#### B Rules Explanation

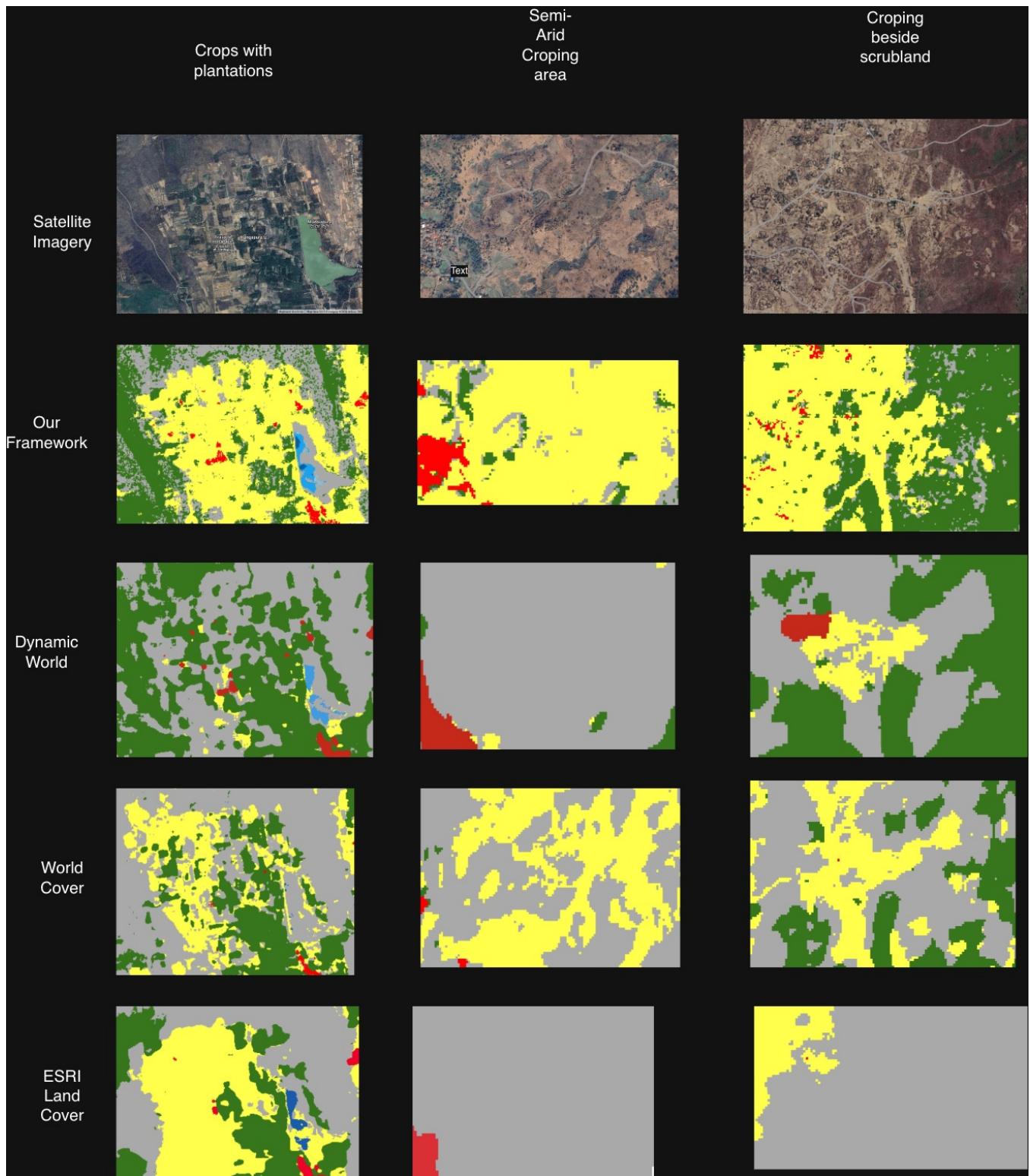
Using these three features together, we empirically defined a set of rules that reliably yields a small number of high-confidence farm and non-agricultural boundaries within each grid. For farmland, we observed that segments usually exhibit entropy values below 1.0, consistent with the relatively uniform interiors of cultivated fields. This threshold also removes high-entropy farm parcels such as plantations or fields with scattered trees or grid-like cultivation patterns; however, since our goal is to obtain only a few clean, high-confidence farm examples rather than an exhaustive set, such

exclusions are acceptable. These more complex fields are expected to be recovered later by the classifier trained on the high-confidence examples.

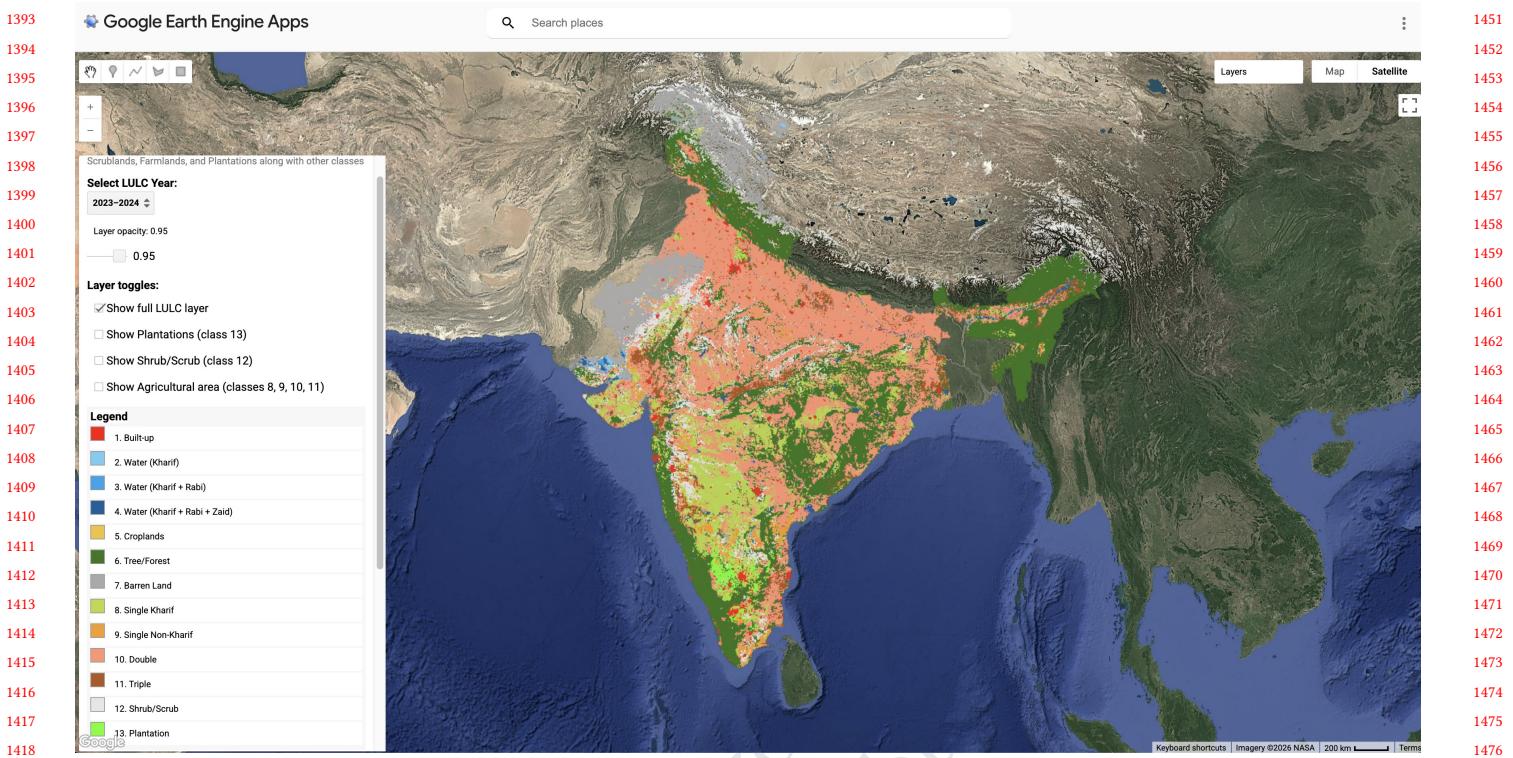
Rectangularity provides a second strong cue: we found that segments with rectangularity values greater than 0.67 give a good balance between admitting mildly irregular farm shapes while filtering out the majority of small non-agricultural artifacts. Finally, we restrict candidate farms to the empirically observed size range of 500–2000 m<sup>2</sup>, which captures most farm parcels in India.

#### C Implementation

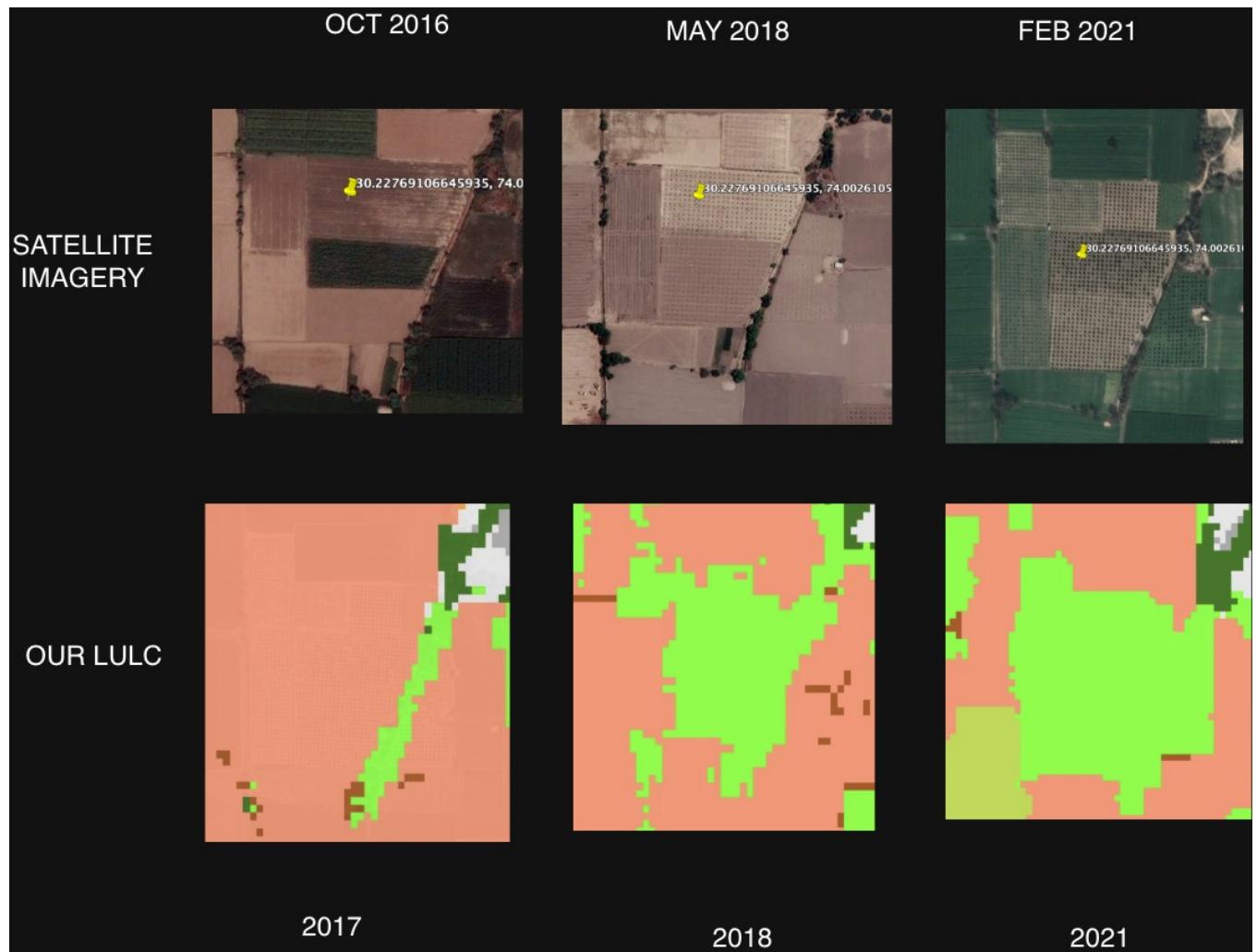
For each subdivision, we record the geographic coordinates (top-left and bottom-right latitude–longitude) in a CSV file to ensure reproducible and consistent downstream processing. These standardized units form the input to the subsequent high-resolution boundary delineation stage.



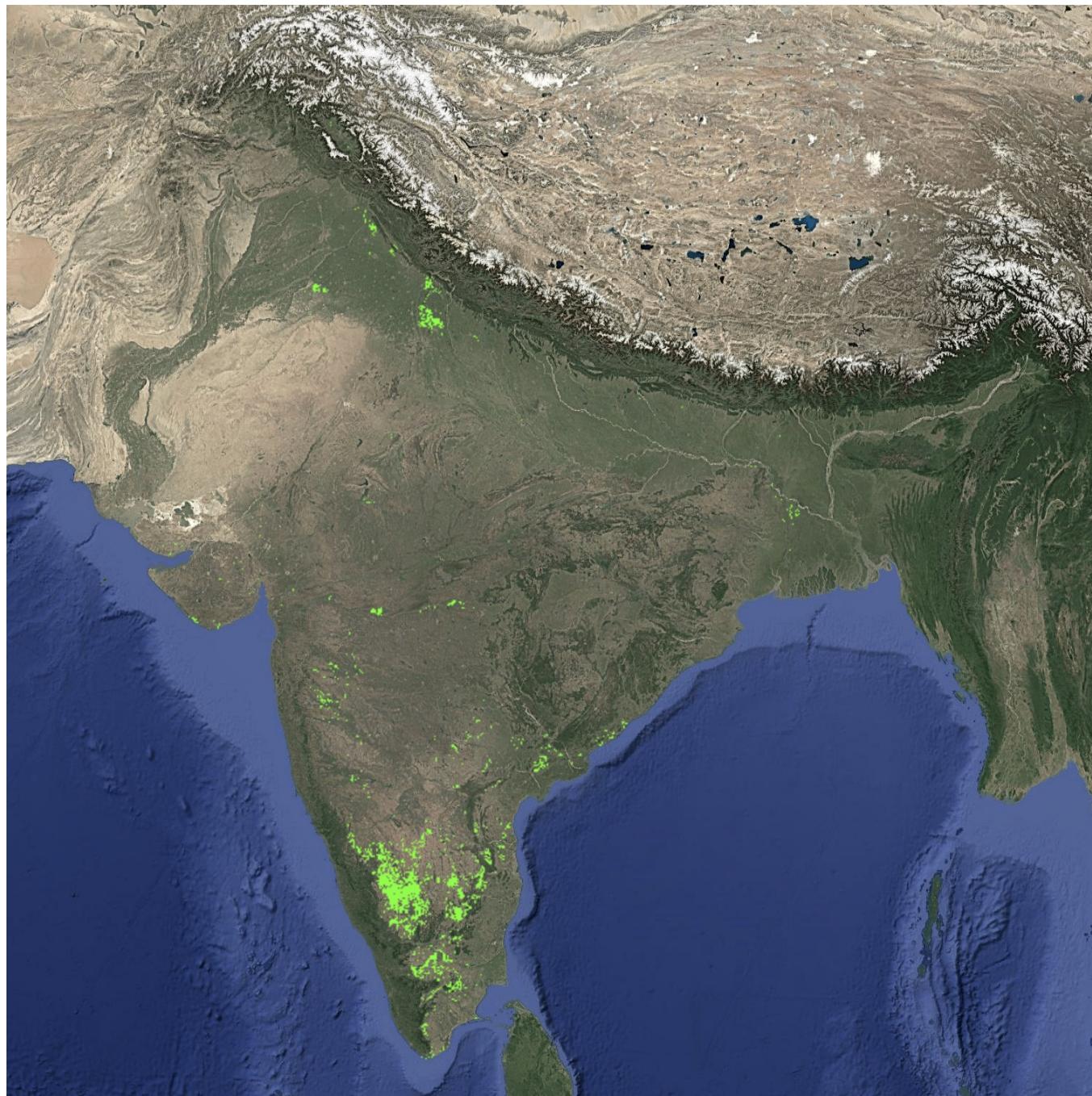
**Figure 7: Outputs of our LULC for various regions. Yellow is agricultural area, grey is scrub and barren, red is urban area and green is tree**



**Figure 8:** Our application showing pan india lulc layer from 2017-2025



**Figure 9: Plantation predictions temporally, Red is double cropping, light green is plantation, dark green is trees, grey is scrub**

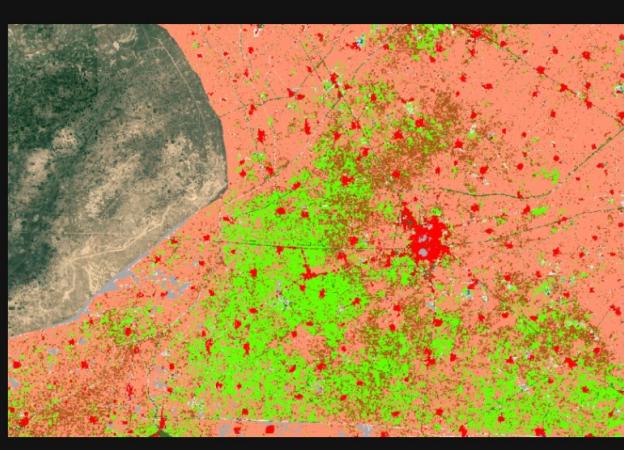


**Figure 10: Pan India Hoticulture (agro plantation) predictions**

1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740



A



B

Figure 11: Kinnow plantation in North India(Abohar)

1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855