

Agricultural Landscape Understanding At Country-Scale

Radhika Dua¹, Nikita Saxena¹, Aditi Agarwal¹, Alex Wilson²,
Gaurav Singh¹, Hoang Tran², Ishan Deshpande¹,
Amandeep Kaur¹, Gaurav Aggarwal¹, Chandan Nath²,
Arnab Basu², Vishal Batchu², Sharath Holla², Bindiya Kurle²,
Olana Missura², Rahul Aggarwal², Shubhika Garg¹, Nishi Shah²,
Avneet Singh², Dinesh Tewari¹, Agata Dondzik², Bharat Adsul³,
Milind Sohoni³, Asim Rama Praveen³, Aaryan Dangi³,
Lisan Kadivar³, E Abhishek³, Niranjan Sudhansu⁴,
Kamlakar Hattekar⁴, Sameer Datar⁴, Musty Krishna Chaithanya⁵,
Anumas Ranjith Reddy⁵, Aashish Kumar⁶,
Betala Laxmi Tirumala⁷, Alok Talekar^{1*}

¹Google DeepMind, India (No contributions to Appendix C).

²Google (No contributions to Appendix C).

³Department of Computer Science IIT Bombay (Only contributions to Appendix C1).

⁴Department of Land Records State Government of Maharashtra (Only contributions to Appendix C1).

⁵TeamUp (Only contributions to Appendix C2).

⁶Department of Agriculture State Government of Telangana (Only contributions to Appendix C2).

⁷Department of Information Technology, Electronics & Communication State Government of Telangana (Only contributions to Appendix C2).

*Corresponding author(s). E-mail(s): atalekar@google.com;

Abstract

Agricultural landscapes are quite complex, especially in the Global South where fields are smaller, and agricultural practices are more varied. In this paper we

report on our progress in digitizing the agricultural landscape (natural and man-made) in our study region of India. We use high resolution imagery and a UNet style segmentation model to generate the first of its kind national-scale multi-class panoptic segmentation output. Through this work we have been able to identify individual fields across 151.7M hectares, and delineating key features such as water resources and vegetation. We share how this output was validated by our team and externally by downstream users, including some sample use cases that can lead to targeted data driven decision making. We believe this dataset will contribute towards digitizing agriculture by generating the foundational baselayer.

1 Introduction

The global food system is facing unprecedented challenges. In 2023, 2.4 billion people experienced moderate to severe food insecurity [1], a crisis precipitated by anthropogenic climate change and evolving dietary preferences. Furthermore, the food system itself significantly contributes to the climate crisis, with food loss and waste accounting for 2.4 gigatonnes of carbon dioxide equivalent emissions per year (GT CO₂e/yr) [2], and the production, mismanagement, and misapplication of agricultural inputs such as fertilizers and manure generating an additional 2.5 GT CO₂e/yr [3]. To sustain a projected global population of 9.6 billion by 2050, the Food and Agriculture Organization (FAO) estimates that food production must increase by at least 60% [1]. However, this also presents an opportunity: transitioning to sustainable agricultural practices can transform the sector from a net source of greenhouse gas emissions to a vital carbon sink. Meeting this escalating demand while mitigating environmental impact necessitates a fundamental paradigm shift in agricultural land management. While significant resources are already allocated to enhance food security globally, their efficacy is constrained by a lack of targeted allocation [4]. This paradigm shift necessitates the digitization of the agricultural sector, enabling precise monitoring and facilitating data-driven decision-making.

The initial step in this digitization is achieving a comprehensive understanding of agricultural landscapes. This is particularly critical in the Global South, where smallholder farms(cultivating less than two hectares of land) are prevalent, and the intricate mosaic of diverse land use, interspersed with vital water resources and natural vegetation, presents unique challenges for accurate mapping and monitoring. These landscapes are characterized by their complexity and dynamism, encompassing features such as cultivated land (e.g., fields, orchards), water resources (e.g., lakes, rivers, canals, ponds, wells), and vegetation (e.g., trees, shrubs, forests). Accurate identification and classification of these features are not only crucial for precise monitoring and targeted resource allocation but also lay the foundation for subsequent layers of digitization. This foundational layer enables the development of sophisticated analytical tools capable of monitoring crop types, spatial distribution, planting dates, and anticipated harvest times.

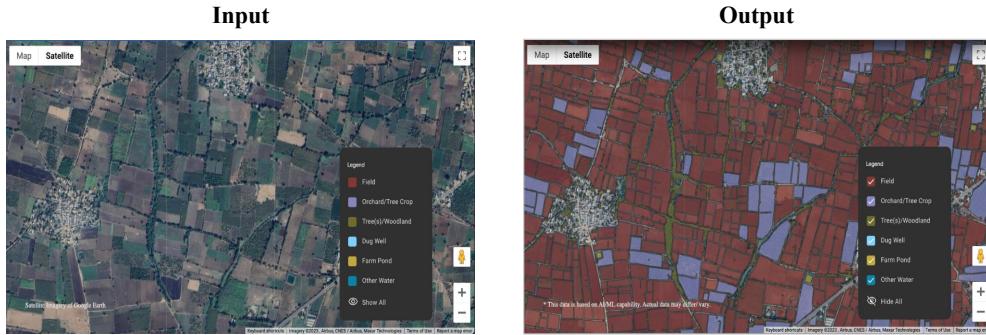


Fig. 1 Demonstration of the model’s ability to extract diverse agricultural features from high-resolution satellite imagery. (Left) Model input: High-resolution satellite imagery. (Right) Model output: Segmented instances of fields, trees, water bodies.

Without this robust understanding of the landscape, constructing accurate and reliable monitoring systems for sustainable agriculture becomes exceedingly challenging. Furthermore, this understanding of agricultural landscapes must encompass the intricate interdependencies within the broader ecosystem. Food systems, water systems, and biodiversity are inextricably linked. Anthropogenic modifications in land use, water management, and agricultural practices can have cascading ramifications on biodiversity, water availability, and ultimately, food security. Therefore, a holistic approach to digitization that captures these interconnections is essential for developing sustainable and resilient agricultural practices.

To achieve this level of granular understanding in a cost-effective and scalable manner, remote sensing technologies are employed. Advances in satellite imagery, with increasingly high resolution and expansive coverage, provide a powerful lens through which to observe and analyze agricultural landscapes at a national scale. This capacity for broad-scale monitoring is essential not only for understanding localized agricultural practices but also for informing regional and national policies related to food security, water resource management, and climate change mitigation. However, extracting meaningful insights from this raw data necessitates sophisticated analytical tools. Machine learning (ML) algorithms, with their capacity for pattern recognition and object classification within images, are ideally suited for analyzing the complex tapestry of features within agricultural landscapes.

We address the challenge of **agricultural landscape understanding (ALU)** by formally modeling it as a multi-class panoptic segmentation problem using the red, green, and blue (RGB) channels of satellite imagery. Panoptic segmentation provides two distinct layers of information: (1) segmenting the image into semantically meaningful regions and (2) identifying individual instances of objects within those regions [5]. Our system outputs segmentation masks, which are pixel-level classifications that delineate the boundaries of key features such as fields, trees, and water bodies, as indicated in Fig. 1.

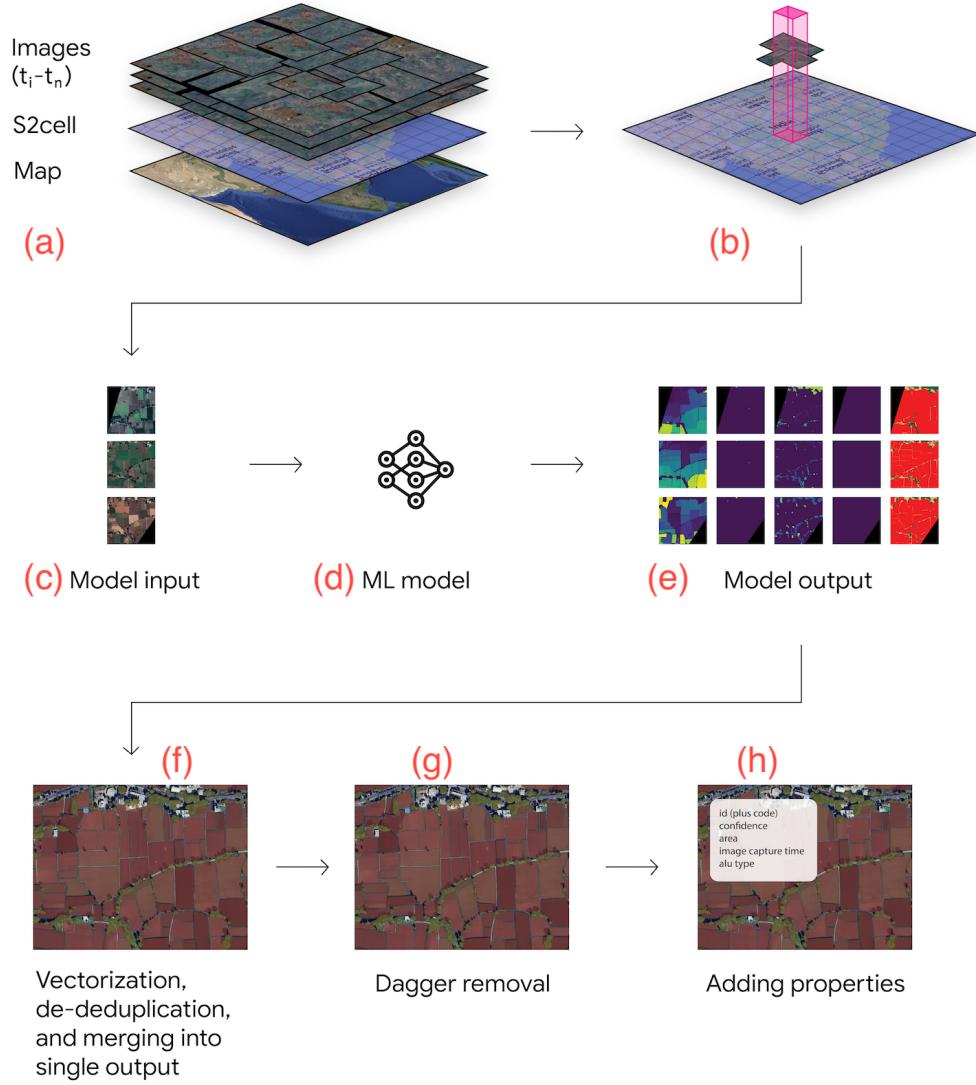


Fig. 2 This figure illustrates the workflow of an ALU system designed to handle the challenges of national-scale agricultural land use mapping. (a) Complete coverage requires multiple satellite images acquired at different times (t_1 to t_n) due to the large study area and satellite orbital constraints. (b) To ensure scalability, the processing is spatially partitioned using S2 cells, a gridded sub-region of the Earth. (c-e) A machine learning model performs multi-class panoptic segmentation on each image to identify agricultural features. (f-h) Post-processing steps, including vectorization, de-duplication, and merging, reconcile information from overlapping images and generate a final output partitioned at s2 cell

1.1 Core Contributions

This study significantly advances the field of agricultural land-use understanding by leveraging high-resolution satellite imagery and machine learning. Our model outputs, accessible via instructions at <http://agri.withgoogle.com>, provide granular insights into agricultural landscapes at a national scale, with a particular focus on smallholder farms. The core contributions of this work are:

National-Scale Mapping with Emphasis on Smallholder Farms This study represents the first comprehensive mapping of agricultural land use at a national scale with sufficient granularity to support targeted interventions, particularly for smallholder farms. We have successfully identified hundreds of millions of smallholder fields and millions of minor irrigation structures, previously unmapped and undocumented. This focus is critical, as smallholder farmers, cultivating less than two hectares of land, comprise 86.2% of India's agricultural workforce and manage approximately half of the total cultivated area [6].

Comprehensive Multi-class Segmentation Moving beyond traditional field boundary delineation, our approach incorporates the identification of trees and water bodies within agricultural landscapes. This multi-class segmentation, novel in the context of satellite-based land-use mapping, provides a nuanced understanding of land-use patterns crucial for diverse applications, including:

Water Infrastructure Assessment Precisely locating ponds and wells enables detailed assessments of water access for individual farms and regions, facilitating the development of sustainable water management strategies.

Agroforestry Analysis Mapping trees within agricultural settings allows for the analysis of agroforestry practices and their potential for carbon sequestration, contributing valuable data to climate mitigation efforts.

Rigorous On-Ground Validation To ensure the reliability and real-world applicability of our model, we conducted meticulous on-ground validation in collaboration with external partners. This validation involved comparing model outputs with in-situ surveys and aligning them with non-geo-referenced data. This rigorous validation process, exceeding the reliance on fixed test sets common in similar studies, strengthens our findings and bolsters confidence in the utility of our approach for informing agricultural interventions and policy decisions.

1.2 Related Work

The analysis of land-use patterns is fundamental to informing resource planning and allocation decisions, particularly in the context of food and water security [7]. Historically, land-use monitoring has progressed through three distinct phases: (a) labor-intensive manual land records [8–10], (b) smartphone-based data collection using tools like Open Data Kit (ODK 1.0) [11] and Ground [12], and (c) remote sensing techniques employing Unmanned Aerial Vehicles (UAVs) and satellites.

Remote Sensing While manual and smartphone-based surveys remain valuable for localized assessments, they are inherently labor-intensive and time-consuming. Remote sensing methods, utilizing UAVs and satellites, have emerged as a more efficient and cost-effective solution for large-scale land-use monitoring. These technologies

capture imagery across extensive areas and a broad spectrum of wavelengths, providing rich data for analysis. While satellite imagery may have lower spatial resolution compared to UAV imagery, its expansive coverage and historical data availability render it particularly well-suited for monitoring agricultural landscapes and informing land management initiatives.

Deep Learning for Remote Sensing Given the complexity and volume of remote sensing data, deep learning approaches have been widely adopted for automated land-use analysis. Convolutional Neural Networks (CNNs), with their capacity to learn hierarchical representations from image data, have become the dominant architecture for segmentation tasks in remote sensing. The U-Net architecture [13], with its symmetric encoder-decoder structure and skip connections, has proven particularly effective in capturing multi-scale context, essential for handling the diverse object sizes encountered in remote sensing images. Our approach also employs a U-Net architecture, detailed in Sec. 2. To address the challenge of limited labeled data in remote sensing, we leverage transfer learning in conjunction with supervised learning to achieve effective semantic segmentation.

Numerous studies have focused on delineating field boundaries from remote sensing data. However, most approaches utilizing publicly available datasets, such as Sentinel-2, operate at a 10-meter resolution [14, 15] which is insufficient for resolving smallholder farms that require sub-meter accuracy. Consequently, this study utilizes high-resolution satellite imagery. While some studies have explored super-resolution techniques [16], alternative satellite sources [17], or data fusion approaches [18] to achieve near sub-meter accuracy, their focus has often been limited to larger fields [19] or specific crop types [20]. Other studies have addressed field delineation across diverse sizes and categories [21–24], but often for a single land-use class. In contrast, this work presents a novel approach by performing multi-class segmentation at high resolution, enabling the identification of not only field boundaries but also trees and water bodies within smallholder farming systems.

Labeled Datasets and Temporal Context The development of labeled datasets for agricultural applications has progressed rapidly, with initiatives such as AI4SmallFarms [25], PASTIS [26], and Pula/Tetra Tech’s Bird’s Eye [27] providing valuable resources. However, these datasets often lack temporal context, limiting their utility for assessing the accuracy of field segmentation in dynamic agricultural environments. Our dataset distinguishes itself by explicitly incorporating timestamps with field segmentation outputs. This unique approach facilitates time-sensitive analysis, which is crucial in regions like India, where agricultural landscapes undergo rapid transformations due to seasonal variations, crop rotations, land-use modifications, and ownership changes. By capturing the temporal dynamics of agricultural fields, our dataset enables a more comprehensive understanding of land-use patterns and empowers stakeholders to make informed decisions.

2 Agricultural Landscape Understanding

2.1 Overview

This study addresses the challenge of Agricultural Landscape Understanding (ALU) by employing multi-class panoptic segmentation on high-resolution satellite imagery. Panoptic segmentation provides a rich representation of agricultural landscapes by classifying each pixel into semantic categories (e.g., fields, water bodies, trees) while simultaneously identifying individual instances within those categories (e.g., distinct fields, individual ponds). This granular understanding is crucial for informing agricultural management practices, resource allocation, and policy decisions.

Fig. 2 illustrates the overall workflow for our ALU system. Due to the expansive geographic extent of the study area (in this case, an entire country) and orbital constraints of satellites, a single satellite image cannot encompass the entire region of interest and multiple images acquired at different times (t_1 to t_n) are required to achieve complete coverage (Fig. 2a). This temporal variation in image acquisition introduces the challenge of reconciling potentially disparate representations of the landscape, as weather patterns and seasonal phenomena (e.g., changes in crop stage) can significantly alter the appearance of agricultural features over time. Our system addresses this by systematically processing each image in the collection through our machine learning model, which performs multi-class panoptic segmentation to identify and delineate various agricultural features (Fig. 2c-e). To reconcile temporal variations, we rely on sufficient overlap between images acquired at different times. The resulting segmentation masks are then subjected to a series of post-processing steps, including vectorization, de-duplication, and merging, to generate a comprehensive and spatially accurate representation of the agricultural landscape, effectively merging information from overlapping images while accounting for temporal inconsistencies (Fig. 2f-h).

To ensure scalability and computational efficiency, the processing workflow is sharded by S2 cells [28] (Fig. 2b). S2 cells are a hierarchical spatial indexing system that subdivides the Earth's surface into a grid of cells, providing a convenient and efficient way to organize and access geospatial data. This spatial partitioning strategy allows for parallelization of the image processing and analysis, enabling the efficient handling of vast amounts of data required for national-scale land-use mapping.

2.2 Dataset Creation

Creating a high-quality annotated dataset is crucial for developing and evaluating machine learning models for agricultural land-use understanding, especially given the lack of standardized benchmarks at high resolution and lack of existing consensus on dataset labeling guidelines, which presents a bottleneck within agricultural remote sensing research. We have built a comprehensive dataset comprising high-resolution satellite imagery and corresponding pixel-level annotations.

2.2.1 Labeling Process

The labeling process involved the manual annotation of 1119 image tiles, each measuring 1500 x 1500 pixels (corresponding to a ground sampling distance of 30 cm and

a spatial extent of 450 m x 450 m). These tiles were randomly sampled from a collection of high-resolution satellite imagery acquired over India between 2019 and 2021. Annotators were tasked with delineating various agricultural features within each tile, classifying them into predefined semantic categories: "ground classes" (farm field, dug well, farm pond, other water) and "overhead classes" (trees/woodland, cloud). Tertiary classes, "unlabelled/other" and "ignore," were used to categorize areas outside the scope of the primary analysis or those with ambiguous features. Fig. 3 provides a visual representation of these semantic classes and their hierarchical organization.

It's important to note that certain classes can intersect with others, while some cannot. For instance, a tree can be located within a field, and clouds can obscure any underlying class. However, a pond cannot intersect with a field, as these represent mutually exclusive land-use types. These relationships between classes were carefully considered during the annotation process to ensure accurate and consistent labeling.

Fig. B1 and Sec. A illustrate the annotation process with an example input image and its corresponding annotation output. As shown in the figure, annotators meticulously outlined the boundaries of different features, such as farm fields, farm ponds, and woodlands, ensuring accurate representation of the diverse elements within the agricultural landscape. This detailed annotation process captures both the semantic categories and the spatial extent of individual instances, providing the necessary information for training and evaluating our panoptic segmentation model.

Detailed annotation guidelines, ensuring consistency and accuracy in labeling, are provided in the supplementary material. Table 1 presents the distribution of labels across different semantic classes. As evident from the table, the dataset exhibits class imbalance, with a significantly lower prevalence of wells and ponds compared to fields and trees. This imbalance reflects the natural distribution of these features in agricultural landscapes.

Features	Labelled Dataset
Fields	105955
Ponds	456
Other Water	1083
Trees	101825
Wells	332

Table 1 Distribution of Features in Labelled Dataset

2.2.2 Dataset Preparation

To facilitate model training and evaluation at different spatial resolutions, we created two versions of the dataset: one with the original 30 cm resolution imagery and another with 1 m resolution imagery obtained through downsampling. Next we performed rasterization, converting the vector annotations into a raster format where each pixel is assigned a label.

To further align the data with the requirements of our model architecture (as explained in the following section), we generated multiple layers of labeled data to cater

to the multi-task learning nature of our model. Specifically, we created five distinct layers: one for semantic segmentation and four for instance segmentation.

Following rasterization, the imagery was divided into smaller patches of 500 x 500 pixels to facilitate efficient processing and memory management during model training. Finally, we applied a conventional train-validation-test split to the processed dataset to ensure robust model evaluation and generalization assessment.

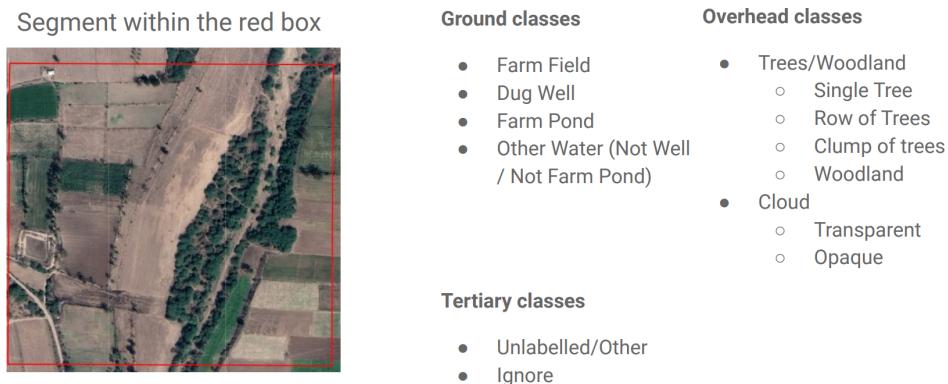


Fig. 3 Annotators were provided guidelines to segment the landscape within the red box (i.e., the tile on the left) into the given categories (right).

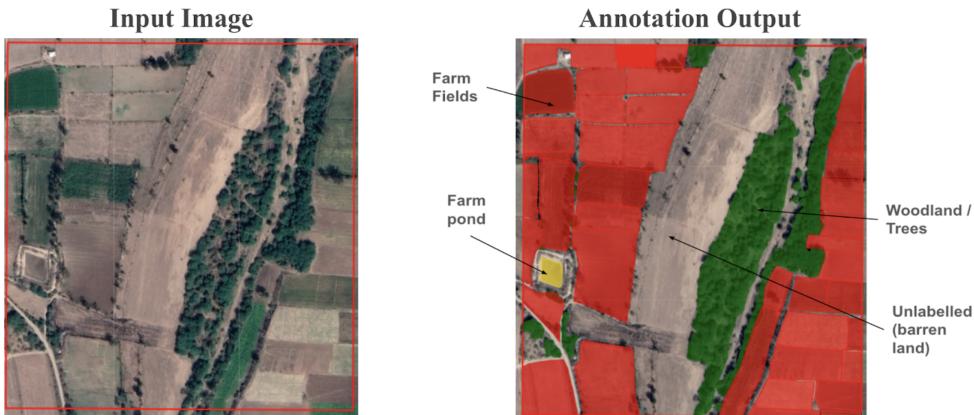


Fig. 4 Example of an annotated sample. Left: input high resolution satellite image. Right: human labeled annotations

2.3 Task Description

As previously mentioned, we pose the overall problem of ALU as that of multi-class panoptic segmentation, where we must identify multiple (potentially overlapping) instances of various classes, along with labeling each instance with the appropriate class. Given an input image, the model must assign class labels and instance labels to each pixel, where an instance is defined by the collection of pixels sharing an instance label. For our problem setting, each pixel can have multiple class labels and instance labels, with some constraints. This is illustrated in Fig. 5

Layer	Classes
Ground	fields, farm ponds, other water bodies
Well	dug wells
Tree	trees/woodland
Cloud	opaque cloud, transparent cloud
Tertiary	background, ignore

Table 2 The dataset comprises four layers: ground, well, tree, and cloud, each containing instances of the corresponding classes. The "background" class denotes areas not falling into any specific layer category, while "ignore" indicates masked regions. It is important to note that instances within a single layer cannot intersect with each other. However, instances across layers can intersect. For example, a field can intersect with a tree but not with another field. Note that throughout this paper and for model training, we focus only on the following four layers, namely ground, well, tree, and cloud. The 'tertiary' layer is detailed for the sake of completeness but is not directly relevant for modeling.

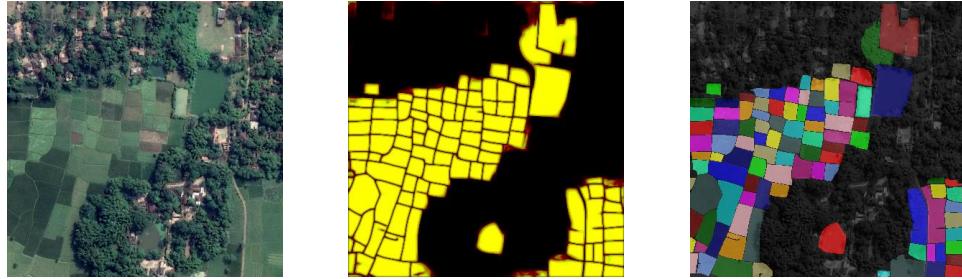


Fig. 5 Visualizing the Task: This figure illustrates our objective: given an input satellite image with RGB bands (left), our goal is to generate both a semantic segmentation map and an instance segmentation map. The center and right images showcase these outputs for the ground instance layer, highlighting the distinction between class-level and object-level segmentation.

Concretely, for each input image of shape (384, 384) which is obtained after the cropping of the (500,500) patches, we must produce a per-pixel output of the set of labels for each pixel of the form (class labels, instance labels).

2.3.1 Classes and Layers

The dataset comprises the following classes: fields, trees/woodland, dug wells, farm ponds, other water bodies, opaque cloud, transparent cloud, background, and ignore. 'Background' denotes any pixel not belonging to the ground, well, tree, or cloud layers, while 'ignore' indicates masked regions excluded from the analysis. We define four distinct layers of entities: ground, well, tree, and cloud (as noted in Table 2). The rationale for establishing these layers is threefold:

Handling Overlap A single pixel can belong to multiple instances across different layers (e.g., a tree on a field). The labeling scheme and model architecture must accommodate this potential overlap.

Computational Efficiency and Physical Constraints By separating features into layers, we enforce physical constraints (e.g., a tree can overlap with a field, but a pond cannot overlap a field) and reduce the computational dimensionality of the problem.

Representing Z-axis The layers conceptually represent the z-axis (height) of landscape features. Features at the same z-axis level (e.g., ground features like fields and ponds) belong to the same layer. Entities within a layer cannot overlap, but entities across layers can. The model must identify overlapping instances while adhering to these layer-specific constraints.

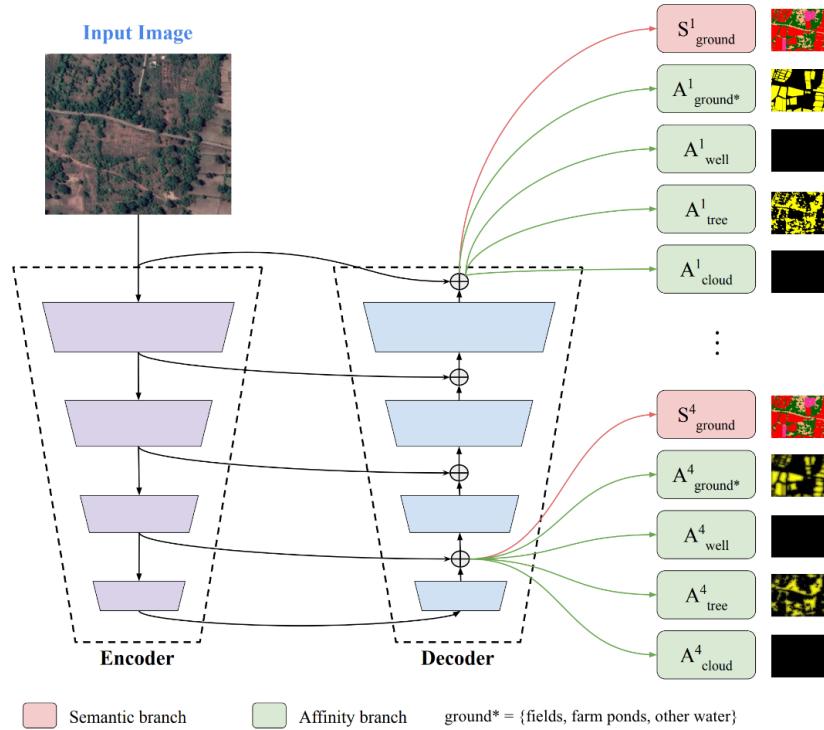


Fig. 6 Our model employs a U-Net style encoder-decoder framework. Model has five distinct output heads: one semantic segmentation head and four affinity mask heads per upsampling stage.

2.4 Machine Learning model

We adopt an affinity mask-based segmentation approach, as applied in [29]. This approach involves predicting pixel-pair affinities to identify instances, followed by assigning class labels based on semantic segmentation.

2.4.1 Model Architecture

Our model architecture, illustrated in Fig. 6, is based on U-Net [30] with a ResNet50 encoder that has been pre-trained on ImageNet. The U-Net architecture enables multi-scale feature extraction, producing outputs at resolutions of $1x$, $\frac{1}{2}x$, $\frac{1}{4}x$, and $\frac{1}{8}x$. At each scale, the model generates a semantic segmentation map (S_i) and four affinity mask layers (A_i layer for each layer). As depicted in Figure 5, the input image is processed by the encoder, and the resulting feature maps are passed to both the semantic and affinity branches. The semantic branch produces the semantic segmentation maps at different scales. The affinity branch generates the affinity masks, which are 256-dimensional vectors per pixel, defining the probability of neighboring pixels belonging to the same instance within a layer. These affinity masks (A_i layer) are then converted to instances using a graph-based method [29], and each instance is assigned a class label based on the corresponding semantic segmentation map. Note that 'ground*' represents the combined set of ground classes in the ground layer. This multi-scale and multi-task learning approach allows the model to capture both semantic and instance-level information at various levels of detail, contributing to a comprehensive understanding of the agricultural landscape.

2.4.2 Training Objective

We use the same loss functions from [29], with a few modifications. Specifically, there are two types of loss functions used:

Semantic segmentation loss The semantic segmentation map at each scale is trained using CE-Focal loss (Cross-Entropy Focal Loss), which combines softmax cross-entropy loss with focal loss [31].

Affinity mask Each affinity mask is trained using instance-aware pixel-pair affinity hierarchically (as shown in Fig. 6). The pixel-pair affinity method computes the probability that two pixels belong to the same instance at multiple scales in a hierarchical manner. Hierarchical manner implies that the affinity is computed at multiple scales, starting with neighboring pixels and progressively considering larger distances. The loss is computed across affinity masks at all resolutions. The mathematical formulation for the loss function is same as [29] with some additional weights to the edge pixel. We add an extra weighting to pixels that lie at the edge of an instance for better boundary delineation. The loss functions are applied to outputs produced at all scales, giving the model feedback at multiple resolutions.

2.4.3 Model Selection Criteria

We use the mean instance-level Intersection over Union (IoU) average over each class to pick the best checkpoint in the training process. We evaluate the model on additional metrics which are discussed in Sec. 3

2.5 Post Processing

The raw output from the model undergoes several post-processing steps to refine the segmentation results and generate the final land-use map. These steps are crucial for ensuring the accuracy, consistency, and usability of the data for downstream applications.

Vectorization The initial model output consists of pixel-based raster segmentation masks. These masks are vectorized to create polygon representations of the identified features. Vectorization provides a more compact and efficient representation of the land-use features, facilitating subsequent processing and analysis.

De-duplication Due to the nature of satellite image acquisition, multiple images may cover the same geographic area, resulting in overlapping and potentially redundant predictions. A de-duplication process is employed to reconcile these overlaps and ensure that each feature is represented uniquely in the final output. This process considers various criteria, such as the coverage area and the recency of the images, to prioritize the most reliable predictions. For the purpose of reconciling and avoiding edge artifacts data across shards needs to be observed holistically, limiting opportunities of parallelizing operations.

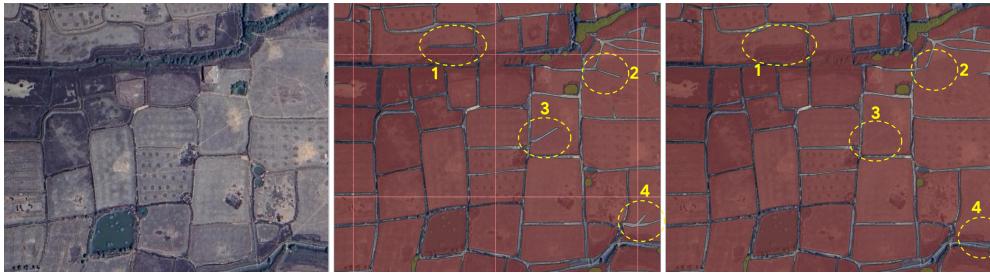


Fig. 7 Left to right: Input satellite image, model output, model output post processed via *Dagger removal* (the sharp protrusions present in the field polygons are smoothed out).

Smoothing and Regularization The vectorized polygons may exhibit irregularities or artifacts due to limitations in the model’s predictions or noise in the input imagery. To address this, a smoothing and regularization step is performed to refine the boundaries of the polygons, resulting in a more visually appealing and accurate representation of the land-use features. This step involves removing sharp angles or “daggers” (as indicated by the dashed polygons in Fig. 7) and applying smoothing algorithms to create more natural and realistic shapes.

Feature Identification To enable efficient indexing and referencing of individual features, an identifier is assigned to each feature based on the Plus Code [32] of its centroid. Plus Codes are a geocoding system that provides short, alphanumeric codes for locations, offering a human-readable and easily shareable way to identify geographic coordinates. By using the Plus Code of the centroid, we can assign a mostly unique identifier to each feature, facilitating efficient data management and retrieval.

Spatial Indexing and Data Partitioning To facilitate efficient data storage, retrieval, and analysis, the processed data is partitioned using S2 cells at level 13. This partitioning strategy allows users to query and retrieve data for specific regions of interest, facilitating targeted analysis and application. A level 13 S2 cell covers an area of approximately 1 km x 1 km and typically requires about 1 MB of storage. This granularity offers a good balance between data volume and spatial coverage, making it suitable for sharing data as API responses while providing sufficient information for various analyses and applications. The final output is stored in GeoJSON format, with S2 cells as the fundamental spatial unit. GeoJSON is an open standard format for representing simple geographical features and their non-spatial attributes using JavaScript Object Notation (JSON). This format is widely used in web-based mapping applications and geographic information systems due to its human-readability and compatibility with various software tools. This sharding strategy allows users to access data for specific regions by querying with S2 cell IDs, geographic coordinates, or region names. This enables flexible and targeted access to the rich information contained within the ALU output, facilitating a wide range of analyses and applications.

Identification and Exclusion of Non-Agricultural Areas To further enhance the utility of the data for agricultural applications, we identify and flag S2 cells that are unlikely to contribute significantly to agricultural production. These include areas classified as urban, desert, and hilly terrain, which typically exhibit low agricultural activity. However, recognizing that many S2 cells may only partially overlap with these non-agricultural regions, we apply a threshold-based exclusion criterion. This criterion ensures that only cells with a substantial portion of their area occupied by non-agricultural land are flagged, thereby preserving cells with mixed land use that may still contain agriculturally relevant features.

3 Model Evaluation

In this section, we first discuss different failures/ errors that may arise in segmentation tasks, followed by quantitative and qualitative evaluations of our model.

3.1 Segmentation Errors and Failure Types

Segmentation models can exhibit a variety of undesirable behaviours that may affect downstream use cases. Some of the important types of failures are:

Under-segmentation This occurs when one predicted instance covers multiple ground truth instances. Fig. 8a. presents an example where a single field is predicted when,in reality, there are multiple fields

Over-segmentation This occurs when one ground truth instance is covered only when multiple predicted instances are combined. Fig. 8b. presents an example where multiple fields are predicted when,in reality, there is a single field.

False negatives This occurs when the model fails to detect an instance. Fig. 8c. presents an example where a farm pond is missing in the segmentation output.

False positives This occurs when the model predicts an instance where, in reality, there is none. Fig. 8d. presents an example where the model predicts a field over a railway track.

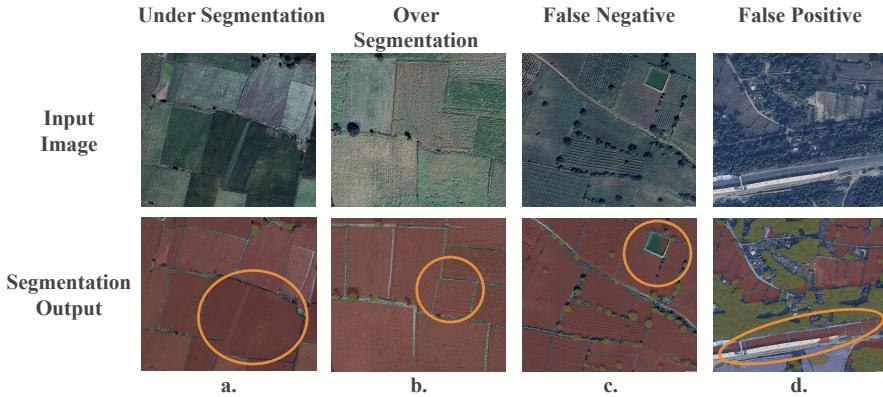


Fig. 8 Visualisation of the different failure modes of a segmentation model (under segmentation: predicting a single instance for multiple ground truth instances, over segmentation: predicting multiple instances in place of a single ground truth, false positive: hallucinating an instance in the prediction, false negative: omitting an instance in the prediction).

Note that these categories are not completely independent, e.g. under-segmentation can naturally give rise to false negatives.

3.2 Metrics and Quantitative Evaluation

Panoptic segmentation models, such as the model we train in this work, are typically evaluated with a metric known as ‘panoptic segmentation quality’ [5]. This metric quantifies the degree of overlap between the predicted instances and the ground truth instances, along with the accuracy of classifying the type of each instance. This provides a single metric to characterize the overall performance of the model and has been widely used in literature [33],[34]. However, a single metric cannot capture all aspects of the model’s behavior. Recognizing that different downstream applications might be sensitive to different kinds of prediction errors, we seek to comprehensively understand the performance of a segmentation model by using multiple metrics, each targeting a distinct failure mode discussed earlier.

The various metrics we use and our model’s performance are discussed below.

We evaluate our model on our labeled data with four metrics:

1. **mIoU (mean Intersection-over-Union)** mIoU provides an aggregated evaluation that encompasses all kinds of errors in segmentation. We use a modified version of mIoU as we find it more suitable for the data we are dealing with. Our mIoU is computed as follows. We first match all predicted instances to ground truth instances per class. A predicted instance is considered to match with a ground truth instance if they both belong to the same class AND the overlap covers at least 10% of the ground truth instance. All predicted instances that match the same ground truth instance are merged together, and predicted instances that do not match any ground truth instance are ignored. We then compute the IoU of the merged predictions w.r.t the matching ground truth instances, and take the average over all ground truth instances per class. We allow multiple predictions to match with a

Class	Mean IOU	Median IOU	Over Seg	Total GT	False Negative Rate	Under Seg	Total PRED	False Positive Rate
Fields	0.68	0.81	1.15	88760	7.09	1.11	132675	34.39
Ponds	0.05	0.00	1.00	163	93.87	1	1118	99.19
Trees	0.37	0.41	1.16	88116	30.2	1.06	151120	60.01
Clouds	0.23	0.03	1.05	211	57.82	1.16	1108	86.91
Wells	0.03	0.00	1.00	26	91.98	1.00	237	38.82

Table 3 Evaluation Metrics for Different Categories

ground truth because, for agricultural landscape understanding we find that it does not hurt any downstream applications. E.g. if one field is segmented into multiple smaller fields then any analysis of a particular area of interest remains unaffected.

2. **False Negative Rate** This is defined as the fraction of ground truth instances that do not have any matching predicted instance (match defined in the same way as previously). This metric is also a standard metric.
3. **Over segmentation Metric** We introduce a metric to quantify the degree of over-segmentation by our model, i.e. how frequently does our model break up one ground truth instance across multiple predicted instances. It is defined as:

$$\text{Over-Seg Metric} = \frac{\sum_{i=1}^N \text{PRED_inst_count_per_GT_inst}_i}{\sum_{i=1}^N \mathbf{1}_{\{\text{PRED_inst_count_per_GT_inst}_i > 0\}}}$$

where $\text{PRED_inst_count_per_GT_inst}_i > 0$ is the total number of predicted instances for which at least one GT instance is present, $\text{PRED_inst_count_per_GT_inst}_i$ is the number of unique overlapping inferred instances per i -th ground truth instance, If $\text{PRED_inst_count_per_GT_inst}_i > 1$, then the instance is over-segmented.

4. **Under segmentation Metric** We introduce a metric to quantify the degree of under-segmentation by our model, i.e. how frequently does our model cover up multiple ground truth instances with a single predicted instance. It is defined as:

$$\text{Under-Seg Metric} = \frac{\sum_{i=1}^N \text{GT_inst_count_per_PRED_inst}_i}{\sum_{i=1}^N \mathbf{1}_{\{\text{GT_inst_count_per_PRED_inst}_i > 0\}}}$$

where $\text{GT_inst_count_per_PRED_inst}_i > 0$ is the total number of predicted instances for which at least one GT instance is present, $\text{GT_inst_count_per_PRED_inst}_i$ is the number of unique overlapping ground truth instances per i -th inferred instance. If $\text{GT_inst_count_per_PRED_inst}_i > 1$, then the instance is under-segmented.

3.2.1 Quantitative Results

As shown in Table 3, our model achieves strong performance on the majority class (Fields exhibited mean IOU = 0.68). However, performance is significantly lower for Ponds & Wells (mean IOU = 0.05 and 0.03 respectively), likely due to their severe under-representation in the training data.

We are unable to baseline these results due to lack of comparable benchmarks available, especially at the same resolution or on the underlying imagery.

3.3 On Ground Validation

Our partners have performed rigorous on ground validation with this dataset.

Across 10 pilot villages in Maharashtra, our partners at IIT Bombay and Department of Land Record (DoLR), Government of Maharashtra, were able to demonstrate use of our dataset in digitizing land records by reconciling dated and non-georeferenced paper maps with our satellite observations in a mostly automated way. The algorithm developed by IIT Bombay based on our dataset and paper maps provided by DoLR generated outputs which were able to satisfy the legal requirements set by DoLR. More details can be found in Sec. C.1

In 19 villages in Telangana, in a partnership with the Department of Agriculture of the state Government of Telangana and a local startup, TeamUp, the field level dataset achieved an accuracy of 82%. More details can be found in Sec. C.2.

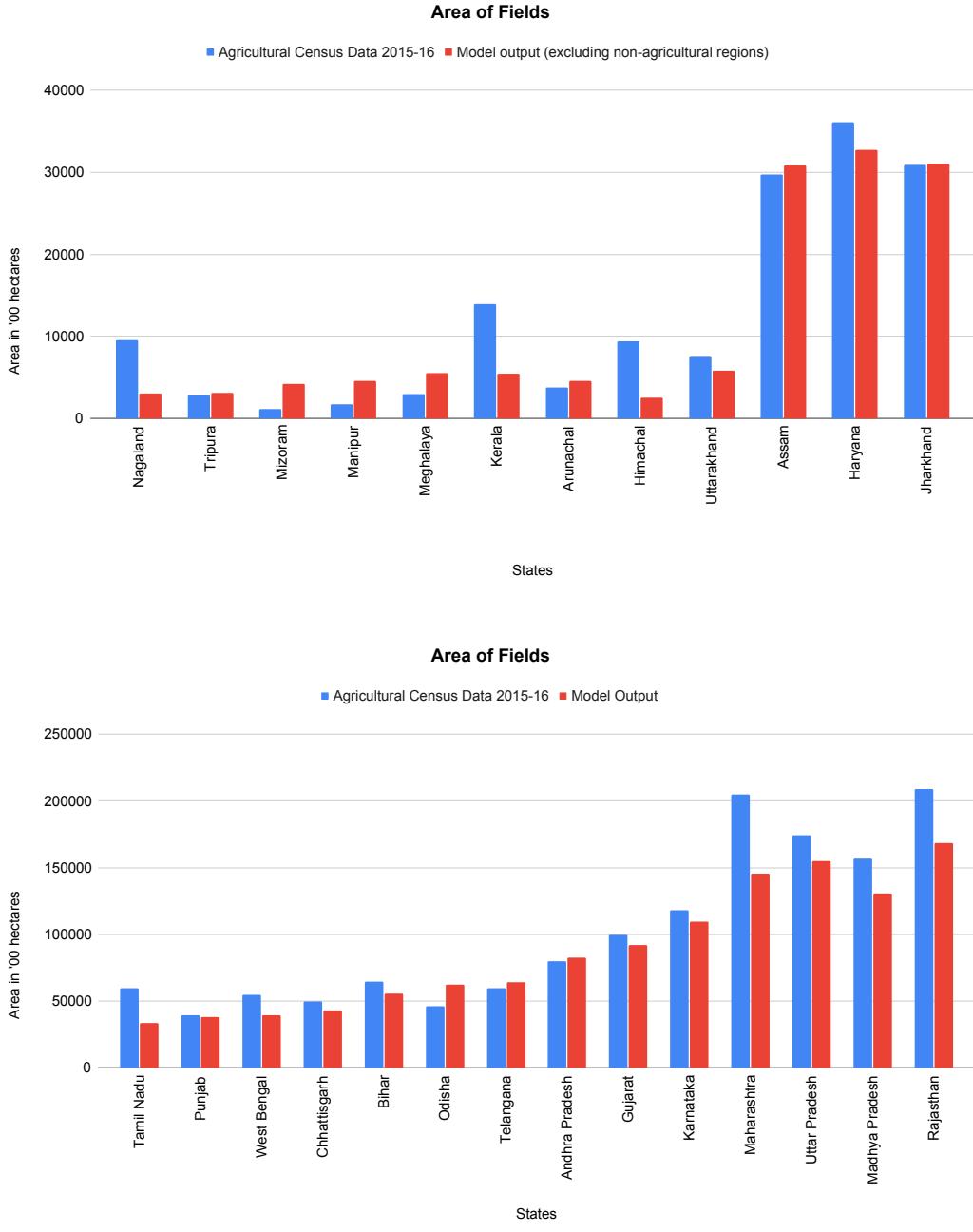


Fig. 9 Comparative analysis between [6] and our model's output, excluding non-agricultural regions. We observe that our model's performance varies with variance in terrain and geographical features. Our model demonstrates comparable performance in regions characterized by plain topography (Bihar, Punjab). However, we observe some divergence in terrains and coastal regions(Kerala, Nagaland, Tamil Nadu, Himachal Pradesh, and Uttarakhand.)

3.4 Aggregate Analysis & Geo-spatial Distribution

Fig. 9 displays a comparative analysis between [6] and model outputs, excluding non-agricultural regions. We observe that our model’s performance varies with variance in terrain and geographical features. Our model demonstrates comparable performance in regions characterized by plain topography, such as Bihar, Gujarat, Jharkhand, and Punjab. However, we observe some divergence in hilly terrains and coastal regions, such as Kerala, Nagaland, Tamil Nadu, Himachal Pradesh, and Uttarakhand.

Detailed tabular data containing state-wise counts and areas of landscape features are present in Sec. B.

4 Conclusion

This work demonstrates the feasibility of mapping agricultural landscapes at a national scale with high granularity, even in smallholder-dominated regions. Our approach, based on multi-class panoptic segmentation of high-resolution satellite imagery, allows for the identification and delineation of diverse agricultural features, including fields, trees, and water bodies. This detailed mapping provides valuable insights into land-use patterns, water resource availability, and agroforestry practices, which are crucial for informing sustainable agricultural interventions and policy decisions. We have also contributed to the research by developing comprehensive annotation guidelines for this domain.

We present high-resolution satellite imagery-derived segmentation outputs for various agricultural landscape features at a national scale, which have been extensively verified for downstream use cases by our partners (See Sec. C). This demonstrates the potential for general landscape understanding at a sub-global scale, with future work aimed at global expansion.

However, our approach has limitations. We were able to incorporate locally relevant landscape features like farm ponds, which are likely less relevant globally. A good general global landscape segmentation model should be able to expand vocabulary to locally relevant features. The class imbalance in the dataset poses challenges for accurately identifying less frequent features like wells and ponds. Lack of locally relevant feature vocabulary and corresponding annotated labels limits our ability to scale and handle distribution shifts with ease, given the degree of supervision needed for the model. Our current approach is also limited by the availability of fresh high-resolution imagery. Landscape features (especially field boundaries) can (and do) change each agricultural season, and an ideal approach is capable of identifying appropriate field boundaries in-season.

Future work will focus on addressing these limitations and further improving the model’s accuracy and generalization capabilities. By harnessing the power of remote sensing and machine learning, we can unlock valuable insights to support sustainable and resilient agricultural systems worldwide.

References

- [1] FAO, U.W. IFAD, WHO: The State of Food Security and Nutrition in the World 2023. FAO, Rome (2023). <https://doi.org/10.4060/cc3017en>
- [2] FAO: FAOSTAT Climate Change: Agrifood systems emissions, Emissions Totals. Accessed: 2024-09-17 (2023). <http://www.fao.org/faostat/en/#data/GT>
- [3] Gao, Y., Cabrera Serrenho, A.: Greenhouse gas emissions from nitrogen fertilizers could be reduced by up to one-fifth of current levels by 2050 with combined interventions. *Nature Food* **4**, 170–178 (2023) <https://doi.org/10.1038/s43016-023-00698-w>
- [4] OECD: Agricultural Policy Monitoring and Evaluation 2023: Adapting Agriculture to Climate Change. OECD Publishing, Paris (2023). <https://doi.org/10.1787/b14de474-en>
- [5] Kirillov, A., He, K., Girshick, R., Rother, C., Dollár, P.: Panoptic Segmentation (2019). <https://arxiv.org/abs/1801.00868>
- [6] Agriculture, C., Farmers Welfare, M.o.A., Farmers Welfare, G.o.I.: All India Report on Agriculture Census 2015-16
- [7] LandRecords. Accessed: 29 August 2024. <https://www.downtoearth.org.in/agriculture/agristack-linking-land-records-with-database-will-exclude-labourers-farmers-warn-experts-7>
- [8] Hodgman, F.: A Manual of Land Surveying comprising an elementary course of practice with instruments and a treatise upon the survey of public and private lands. Accessed: 17 September 2024. <https://archive.org/details/manualoflandsurv00hodgrich/page/n7/mode/2up>
- [9] AssamLandRecords. Accessed: 29 August 2024. https://landrevenue.assam.gov.in/sites/default/files/swf_utility_folder/departments/revenue_com_oid_6/menu/document/assam_land_record_manual-compressed.pdf
- [10] MHLandRecords. Accessed: 29 August 2024. <https://dspace.gipe.ac.in/xmlui/bitstream/handle/1/38474/GIPE-159571.pdf?sequence=3&isAllowed=true>
- [11] Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., Borriello, G.: Open data kit: tools to build information services for developing regions. In: Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development. ICTD '10. Association for Computing Machinery, New York, NY, USA (2010). <https://doi.org/10.1145/2369220.2369236>
- [12] Ground. Accessed: 29 August 2024. <https://groundplatform.org/>

- [13] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. CoRR **abs/1505.04597** (2015) [1505.04597](https://arxiv.org/abs/1505.04597)
- [14] Kaur, A.: Field Boundary Delineation. <https://doi.org/10.13140/RG.2.2.19642.57282>
- [15] Kuang, X.-F., Guo, J., Wang, H.-Y., Wang, H.: Agricultural field boundary delineation using a cascaded deep network model from polarized sar and multispectral images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **16**, 7228–7247 (2023) <https://doi.org/10.1109/JSTARS.2023.3301158>
- [16] Masoud, K.M., Persello, C., Tolpekin, V.A.: Delineation of agricultural field boundaries from sentinel-2 images using a novel super-resolution contour detector based on fully convolutional networks. Remote Sensing **12**(1) (2020) <https://doi.org/10.3390/rs12010059>
- [17] Chaudhury, B., Sahadevan, A.S., Mitra, P.: Agricultural field boundary delineation from multi-temporal irs p-6 liss iv images using multi-task learning. In: 2023 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS), vol. 1, pp. 1–4 (2023). <https://doi.org/10.1109/MIGARS57353.2023.10064490>
- [18] Cai, Z., Hu, Q., Zhang, X., Yang, J., Wei, H., Wang, J., Zeng, Y., Yin, G., Li, W., You, L., Xu, B., Shi, Z.: Improving agricultural field parcel delineation with a dual branch spatiotemporal fusion network by integrating multimodal satellite data. ISPRS Journal of Photogrammetry and Remote Sensing **205**, 34–49 (2023) <https://doi.org/10.1016/j.isprsjprs.2023.09.021>
- [19] Jong, M., Guan, K., Wang, S., Huang, Y., Peng, B.: Improving field boundary delineation in resunets via adversarial deep learning. International Journal of Applied Earth Observation and Geoinformation **112**, 102877 (2022) <https://doi.org/10.1016/j.jag.2022.102877>
- [20] Yang, R., Ahmed, Z., Schulthess, U., Kamal, M., Rai, R.: Detecting functional field units from satellite images in smallholder farming systems using a deep learning based computer vision approach: A case study from bangladesh. Remote Sensing Applications Society and Environment **20** (2020) <https://doi.org/10.1016/j.rsase.2020.100413>
- [21] Wang, S., Waldner, F., Lobell, D.B.: Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision. Remote Sensing **14**(22) (2022) <https://doi.org/10.3390/rs14225738>
- [22] Persello, C., Tolpekin, V.A., Bergado, J.R., de By, R.A.: Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. Remote Sensing of Environment **231**,

- 111253 (2019) <https://doi.org/10.1016/j.rse.2019.111253>
- [23] Long, J., Li, M., Wang, X., Stein, A.: Delineation of agricultural fields using multi-task bsinet from high-resolution satellite images. International Journal of Applied Earth Observation and Geoinformation **112**, 102871 (2022) <https://doi.org/10.1016/j.jag.2022.102871>
- [24] North, H., Pairman, D., Belliss, S.E.: Boundary delineation of agricultural fields in multitemporal satellite imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **PP**, 1–15 (2018) <https://doi.org/10.1109/JSTARS.2018.2884513>
- [25] Persello, C., 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**, 1–5 (2023) <https://doi.org/10.1109/LGRS.2023.3323095>
- [26] Sainte Fare Garnot, V., Landrieu, L., Chehata, N.: Multi-modal temporal attention models for crop mapping from satellite time series. ISPRS Journal of Photogrammetry and Remote Sensing (2022) <https://doi.org/10.1016/j.isprsjprs.2022.03.012>
- [27] Pula/Tetra Tech's Bird's Eye. Accessed: 29 August 2024. <https://ecaas.pula.io/>
- [28] S2Geometry. Accessed: 16 August 2024. <http://s2geometry.io/>
- [29] Gao, N., Shan, Y., Wang, Y., Zhao, X., Yu, Y., Yang, M., Huang, K.: SSAP: single-shot instance segmentation with affinity pyramid. CoRR **abs/1909.01616** (2019) [1909.01616](https://arxiv.org/abs/1909.01616)
- [30] Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation (2015). <https://arxiv.org/abs/1505.04597>
- [31] Lin, T., Goyal, P., Girshick, R.B., He, K., Dollár, P.: Focal loss for dense object detection. CoRR **abs/1708.02002** (2017) [1708.02002](https://arxiv.org/abs/1708.02002)
- [32] Plus Code. Accessed: 16 August 2024. <https://maps.google.com/pluscodes/>
- [33] Cui, Z., Yao, J., Zeng, L., Yang, J., Liu, W., Wang, X.: LKCell: Efficient Cell Nuclei Instance Segmentation with Large Convolution Kernels (2024). <https://arxiv.org/abs/2407.18054>
- [34] Lee, S., Seo, J., Han, K., Choi, M., Im, S.: Context-Aware Video Instance Segmentation (2024). <https://arxiv.org/abs/2407.03010>

Appendix A Annotation Guidelines

To facilitate accurate and consistent labeling of agricultural features by annotators, we established a comprehensive annotation protocol (see Figure B1). Recognizing the variability in class definitions across different implementations, we aimed to create a standardized benchmark for the following classes we annotated. These guidelines were validated by agricultural experts and partners.

For ambiguous cases, such as differentiating barren farm ponds from fields, we provided detailed morphological criteria. For instance, a 'farm pond' was defined as exhibiting clear depth along its edges, even when dry. Similarly, 'wells' were distinguished from 'tree shadows' based on their regular, often circular shape. Annotators were instructed to encompass entire clusters when labeling features like 'trees,' rather than individual instances, to capture the spatial extent of these features accurately. Additionally, we permitted overlapping polygons to reflect the potential overlap of ground, tree, and well layers in the real world.

This annotation protocol serves several purposes beneficial to the community:

- Standardization: It provides a common framework for defining and labeling agricultural features, promoting consistency across different research projects and datasets.
- Reproducibility: The explicit guidelines enable other researchers to replicate our annotation process, enhancing the reproducibility of our findings.
- Benchmark: The resulting dataset can serve as a benchmark for evaluating the performance of machine learning models in agricultural landscape classification tasks.

To ensure annotation quality, we ran pilots to train and calibrate annotators. We also requested annotators to proceed to the next image after spending one hour on an image.

Appendix B Aggregate Analysis & Geo-spatial Distribution

In Table B1 and Table B2, we provide quantitative results as counts of different features and their respective areas. Certain features like trees are not counted because we only identify separated clusters of trees and not individual trees. Table B3 and Table B4 represent the counts of different features and respective areas after excluding non agricultural regions such as mountains, desserts and urban land use.

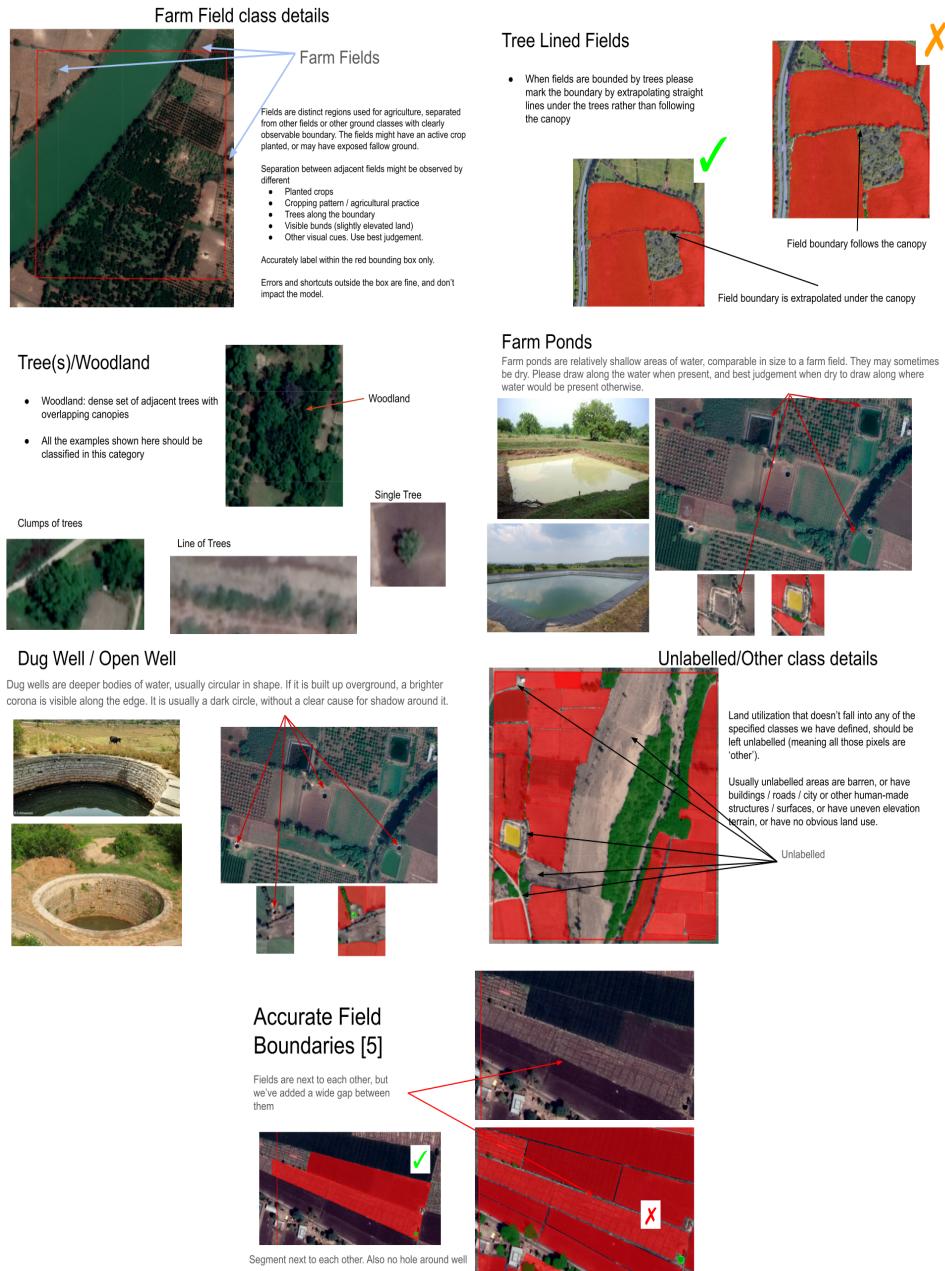


Fig. B1 Annotation Guidelines

Table B1 Quantitative results obtained for counts of different features across India.

State/Region	Dug Well	Farm Pond	Fields
Andaman and Nicobar Islands	5	99	610505
Andhra Pradesh	4881	11076	42231190
Arunachal Pradesh	140	963	9927940
Assam	1003	18115	39791403
Bihar	4764	15422	82691431
Chandigarh	1	13	23155
Chhattisgarh	2166	17810	58422309
Dadra and Nagar Haveli	4	27	357264
Delhi	23	174	375056
Goa	5	84	411781
Gujarat	12932	8662	41578756
Haryana	5275	4521	12868334
Himachal Pradesh	230	940	7917657
Jammu and Kashmir	396	1010	10925429
Jharkhand	3217	13174	62310915
Karnataka	9672	21640	48939810
Kerala	66	2388	5585657
Ladakh	332	466	4995463
Madhya Pradesh	45823	12702	95546315
Maharashtra	52815	37136	72655692
Manipur	537	2355	3942745
Meghalaya	51	658	3741130
Mizoram	4	165	2441003
Nagaland	9	192	1907934
Odisha	1395	30408	79586175
Puducherry	922	1082	2689676
Punjab	4021	5026	15217229
Rajasthan	38120	23380	71311058
Sikkim	31	31	464795
Tamil Nadu	10268	9290	26124579
Telangana	5449	5281	44303937
Tripura	27	2599	3064906
Uttar Pradesh	6453	23698	127111119
Uttarakhand	158	1053	8639589
West Bengal	2776	40191	63542010
Total	213971	311831	1052253947

Table B2 Quantitative results obtained for area covered by different features across India.

State/Region	Farm Pond	Field	Other Water Sources	Trees
Andaman and Nicobar Islands	12.8	54986.32	4082.14	66953.74
Andhra Pradesh	3679.89	8318139.95	147164.39	1709252.35
Arunachal Pradesh	90.9	899889.18	324064.19	1007395.52
Assam	3914.87	3316451.52	100478.7	1299360.04
Bihar	3111.33	5776476.16	43253.75	926699.46
Chandigarh	0.97	2174.34	21.15	2516.78
Chhattisgarh	6837.82	5479545.19	81500.49	1954251.84
DNHDD [*]	3.54	21764.63	337.65	17104
Delhi	26.32	51032.13	1098.66	22975.48
Goa	19	66716.9	3331.04	69122.62
Gujarat	2233.06	9883195.81	102040.2	1335369.47
Haryana	954.09	3363724.93	16111.06	212459.91
Himachal Pradesh	91.41	921780.46	170801.46	758521.67
Jammu and Kashmir	224.94	966021.05	99762.67	675569.9
Jharkhand	2675.52	3371667.18	33490.69	1253515.24
Karnataka	2815.29	11065729.32	85628.08	2002997.94
Kerala	509.31	837844.21	25019.22	826769.61
Ladakh	47.68	753287.64	94485.33	102440.62
Madhya Pradesh	2590.2	17467564.59	113184.09	2540341.96
Maharashtra	4904.44	14694840.41	111663.99	2052305.32
Manipur	711.1	478140.07	51180.31	369553.24
Meghalaya	102.81	556933.36	25992.29	438908.71
Mizoram	23.04	419989.85	39998.25	392882.89
Nagaland	19.34	308434.32	57357.59	308622.21
Odisha	7268.7	6314261.63	96954.29	2570221.85
Puducherry	196.87	331893.77	4819.4	101389.18
Punjab	836.05	3864152.35	12262.58	214762.7
Rajasthan	2734.66	20386083.94	133202.54	1803262.58
Sikkim	2.89	62891.28	30141.01	94230.63
TamilNadu	1658.51	3452558.07	49849.51	1060966.2
Telangana	941.16	6454266.06	72731.38	1118550.91
Tripura	593.35	325045.98	17454.45	272359
Uttarakhand	138.73	1012486.04	174620.49	865189.97
Uttar Pradesh	6355.97	16309133.03	86318.82	2053228.93
West Bengal	10046.13	4173252.29	141015.45	1137016.81
Total	66372.69	151762354	2551417.31	31637069.28

*Area in hectares

[^]Dadar and Nagar Haveli and Daman and Diu

Table B3 Quantitative results obtained for counts of different features across India after excluding s2 cells containing non-agricultural regions

State/UT	Dug Well	Farm Pond	Fields
Andaman and Nicobar Islands	5	90	599040
Andhra Pradesh	4837	10679	41471069
Arunachal Pradesh	66	673	4181101
Assam	885	16122	36495391
Bihar	4553	14131	79263561
Chandigarh	0	5	8911
Chhattisgarh	1365	14289	41601004
Dadra and Nagar Haveli and Daman and Diu	4	22	297609
Delhi	18	64	169898
Goa	4	55	313207
Gujarat	12221	7662	36127685
Haryana	5220	3885	12177978
Himachal Pradesh	54	134	2868517
Jammu and Kashmir	352	729	8544157
Jharkhand	3012	12250	57242032
Karnataka	9594	20957	47917696
Kerala	32	1272	3228771
Ladakh	259	278	3671040
Madhya Pradesh	41956	7320	55501759
Maharashtra	52640	36253	71243004
Manipur	478	1999	3607950
Meghalaya	51	616	3690225
Mizoram	4	163	2434012
Nagaland	7	176	1860035
Odisha	1384	29962	78736036
Puducherry	885	810	2444233
Punjab	3970	4441	14496499
Rajasthan	27298	15149	57228959
Sikkim	0	0	27
Tamil Nadu	10112	8129	24885950
Telangana	5409	5003	43746991
Tripura	22	2357	2854560
Uttar Pradesh	5753	21645	115521177
Uttarakhand	50	449	4719959
West Bengal	2637	36068	59772888
Total	195137	273837	918922931

Table B4 Quantitative results obtained for area covered by different features across India after excluding s2_cells containing non-agricultural regions

State/Region	Farm Pond	Field	Other Water Sources	Trees
Andaman and Nicobar Islands	12.1	54193.69	4051.513944	4051.51
Andhra Pradesh	3609.34	8237729.18	144598.8738	144598.87
Arunachal Pradesh	62.21	457052.55	154191.1538	154191.15
Assam	3522.4	3085859.53	91788.51384	91788.51
Bihar	2923.77	5561620.71	39366.62074	39366.62
Chandigarh	0.51	1141.95	9.555058933	9.56
Chhattisgarh	5879.29	4313268.51	68512.52316	68512.52
DNHDD ^a	2.68	17707.05	277.4700534	277.47
Delhi	16.26	34507.84	447.562889	447.56
Goa	14.1	51530.93	2629.191031	2629.19
Gujarat	2101.21	9213941.48	83790.24241	83790.24
Haryana	879.78	3276586.67	14335.22939	14335.23
Himachal Pradesh	16.31	250811.61	35377.63627	35377.64
Jammu and Kashmir	75.74	746680.59	66165.73885	66165.74
Jharkhand	2510.17	3109014.7	30740.00575	30740.01
Karnataka	2759.6	10966240.07	82948.46262	82948.46
Kerala	317.4	544622.13	15917.53773	15917.54
Ladakh	32.48	525056.77	72157.09529	72157.1
Madhya Pradesh	1413.89	13080637.04	76826.10152	76826.1
Maharashtra	4796.95	14561341.76	108654.1329	108654.13
Manipur	641.51	458305.9	50641.31185	50641.31
Meghalaya	98.45	552655.06	25712.93647	25712.94
Mizoram	22.82	419209.66	39797.46405	39797.46
Nagaland	17.78	304757.04	57183.1892	57183.19
Odisha	7171.74	6243822.79	95012.28819	95012.29
Puducherry	167.36	310013.65	3536.749571	3536.75
Punjab	764.58	3790176.11	11219.14727	11219.15
Rajasthan	1855.76	16841835.52	94428.67964	94428.68
Sikkim	0	2.87	42.68972016	42.69
Tamil Nadu	1494.76	3337609.52	45032.5462	45032.55
Telangana	919.27	6408580.25	70765.1285	70765.13
Tripura	539.13	310917.95	16319.16583	16319.17
Uttarakhand	69.53	580093.3	64165.64959	64165.65
Uttar Pradesh	5951.31	15487152.57	75032.75646	75032.76
West Bengal	9150.96	3924498.79	113277.8988	113277.9
Total	59811.15	137059175.7	1854952.762	1854952.77

*Area in hectares

^aDadar and Nagar Haveli and Daman and Diu

Appendix C External Evaluations and Applications

C.1 Modernizing Land Records in Maharashtra

C.1.1 Introduction

Land records play a crucial role in maintaining property ownership and land use information, with implications for legal, economic, and societal activities. The transition from traditional paper-based records to digital solutions using Geographic Information System (GIS) technology is an important step and has led to great efficiencies, especially in efficient dispute management and land transfer. This appendix describes a pilot project done jointly by IIT Bombay and the Department of Land Records, Govt. of Maharashtra. It utilizes Google ALU outputs, hereby termed as “Google farm plots”, as a key input in the overall modernization of land records in the state of Maharashtra within India.

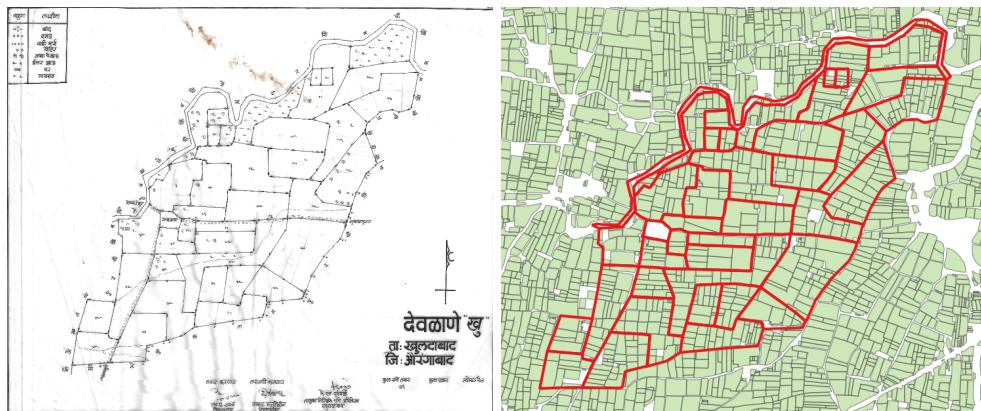


Fig. C2 Modernizing Land Records: Introducing the Problem

C.1.2 Geo-referencing and Data Mismatch

Maharashtra, with an area of about 300,000 sq.km. has roughly 44,000 villages, each with an area between 5-20 sq. km. Each village consists of about 50-300 survey plots. The maps for these villages began as paper or cloth maps stored at the regional land records office. These were eventually digitized and stored as a drawing without geo-referencing (See Figure C2).

An important first step was geo-referencing, i.e., the process of importing these drawings into a GIS system, and subsequently, scaling and translating these maps to align with ground reality, to obtain, what we call as Survey Maps M^0 . This was achieved by the state agencies over the last decade. However, significant mismatches were often observed between the survey maps and agricultural field boundaries (as visible in Figure C3). This is due to many reasons - limitations of historical tools, changes in

land use and ownership and emergence of new infrastructure. This disparity hinders the legal validity and effective use of these digitized maps for various applications. Modification of these survey maps to improve match with agricultural boundaries, without a significant compromise on legal area and original shape, was deemed as a necessary first step to their modernization. This section describes a mathematical and algorithmic formulation of the above problem.

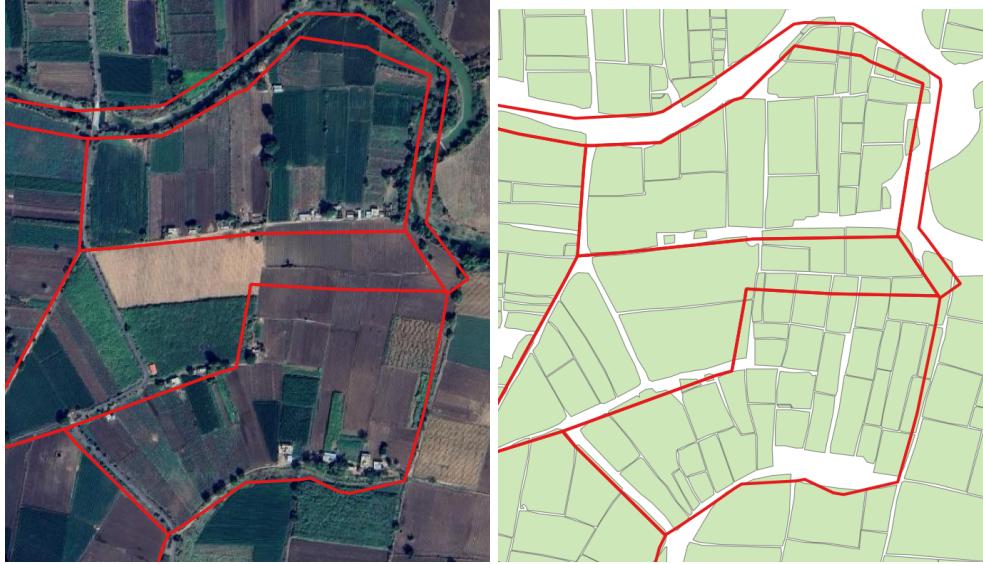


Fig. C3 Mismatch Observed in the input survey maps.

Besides the above mentioned approximate survey maps, an important input was Google farm plots, obtained as collection of geo-referenced polygons that reflect current agricultural boundaries. These polygons provide a discrete view of ground reality, which helps in the formulation of the matching problem as a well-defined optimization problem and devising an algorithm (called **Matchfit**) for the same.

C.1.3 Mathematical Formulation: Two Optimization Problems

Matchfit takes a collection of survey plots as a set of polygons $M = \{Q_1, \dots, Q_k\}$, and the Google farm plots for the region as a collection of polygons $F = \{P_1, \dots, P_s\}$, as the inputs. It is designed as two sequential optimization problems on finding suitable geometric transformations to minimize two related error functions.

We now come to the first error function $ea(M, F)$ which quantifies the mismatch between the family M and F . This is an aggregate of errors $ea(Q, F)$ for individual survey plots $Q \in M$ and F . Heuristically, it is clear that all farm area seen to protrude into a survey plot, or extrude outside, constitutes an error (i.e., a mismatch which is to be minimized). To formalize this notion, we define the **excess area** $ea(Q, F)$ as

follows. For a survey plot Q , and all farm plots $P_i \in F$ which intersects with Q , let the excess area $\text{ea}(Q, P_i) = \min(\text{area}(Q \cap P_i), \text{area}(P_i - Q))$. The excess area $\text{ea}(Q, F)$ of plot Q is the sum of $\text{ea}(Q, P_i)$ over all such intersection P_i . The error function $\text{ea}(M, F)$ is simply the sum $\sum_{Q \in M} \text{ea}(Q, F)$.

The second error function is called the **distance to boundary** or simply $dtb(Q, F)$. This is simply $\text{ea}(Q, F)$ divided by the perimeter of Q , but suitably divided across the edges of Q . The quantity $dtb(Q, F)$ is an estimate of the mean distance by which the boundaries of Q need to be moved to match with the farming area. This plot-specific quantity is a measure of the administrative and legal difficulties in reconciling the survey plot boundary with actual use. Refer to Figure C5 to understand the metric and how it changes.

C.1.4 Description of *MatchFit*

MatchFit proceeds in two steps, called **JitterFit** and **SplineFit**. **JitterFit** takes the original survey map M^0 , and F as input and uses the group T of all translations, rotations and scaling as the allowed transformations. For any $t \in T$, let M^t denote the survey plots obtained by applying t to M^0 . **JitterFit** minimizes $\text{ea}(M^t, F)$ to obtain the map $M^1 = \{Q_1^1, \dots, Q_k^1\}$. This map is thus an optimal geo-referencing of the input map M^0 . Note that scaling is also allowed since the original paper maps did not record aggregate distances.

The second optimization is more tricky and is motivated by the field observation provided to us by the agency. Most villages have regions of good and bad match and the identification of such regions would help in the eventual reconciliation. **SplineFit** is designed around this observation. It proceeds in three steps.

1. *JitterFit* each polygon $Q^1 \in M^1$ independently to minimize the dtb -metric.
2. Select Anchor Plots $A \subset M$.
3. Compute interpolating spline-fit transformation and apply to M^1 to obtain final map M^2 .

In the first step, individual polygons Q_i^1 of M^1 , are unhooked and *JitterFitted* to minimize $dtb((Q_i^1)^t, F)$ to obtain Q'_i . The transformations allowed is the same set T above.

Next, of the set $\{Q'_1, \dots, Q'_k\}$, an index set I is chosen so that the set $A = M'_I = \{Q'_i | i \in I\}$ are those which best fitting and which are most widely distributed across the village. This set is called the **anchor polygon set**.

Finally, a 2D non-linear function $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is found which (i) fixes the anchor polygons and (ii) which is within an ϵ -neighborhood of the identity function. The ϵ is chosen so that distortion in area or shape does not exceed 3%. The final map set M^2 is precisely $f_A(M^1)$.

C.1.5 Implementation and Analysis

The proposed pipeline was successfully implemented in 128 villages, across 5 Talukas in districts across various geographical areas. To verify the accuracy and utility of the algorithm, the generated maps were evaluated on the dtb metric for each plot. Figure C4 shows heatmaps of the dtb metric for the input map M^0 , intermediate map M^1 and output map M^2 for a sample village and Figure C5 shows a close up of a part of the village for a better visualization of the changes made by the algorithm.

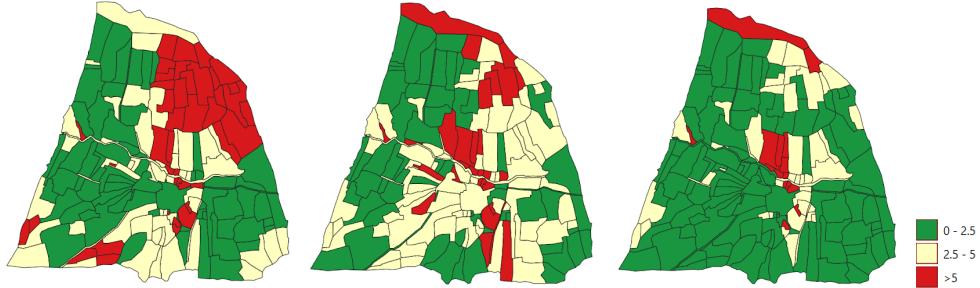


Fig. C4 Distance to Boundary Heatmaps, Input Map M^0 (left), Output Map M^1 (middle) and Output Map M^2 (right) with Green (< 2.5 m), Yellow (< 5 m), Red (> 5 m)

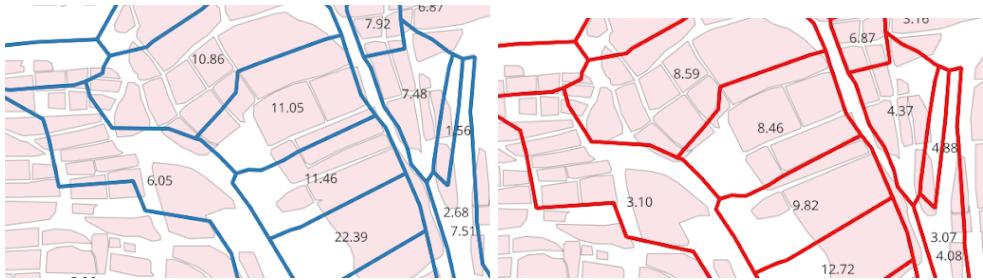


Fig. C5 Distance to Boundary, Before (left) vs After (right)

As shown by the heatmaps and the table below, it is clear that the *MatchFit* algorithm effectively reduces the average dtb . Given that the area and shape distortion is minimal, it is expected that a dtb less than 5m can be resolved by individual farmers without involvement of the department. This helps in reducing the number of cases which need their intervention.

In general, for most of the villages, nearly 60% of the survey plots in the output maps M^2 (also called base maps) have less than 5m distance to boundary which is a substantial improvement over the input maps (see Table C6 for a sample). Moreover, for those plots with higher dtb , the base maps indicate the particular farm plots

village	No. of survey plots	No. of anchor polygons	% of survey plots in M0 with average bund distance < 2.5m	% of survey plots in M0 with average bund distance < 5m	% of survey plots in M2 with average bund distance < 2.5m	% of survey plots in M2 with average bund distance < 5m
akoli	92	17	5.43	41.3	34.44	70
shipgaon	58	12	20.68	41.37	55.56	72.22
kanfodi	66	14	6.06	15.15	30.19	66.04
jasapur	54	6	16.66	46.29	17.65	50.98
indapur	41	9	26.82	63.41	51.43	80
harangul	64	14	15.62	59.37	45.9	72.13
bhaddarpur	63	13	7.93	31.74	41.67	66.67
waghoda	39	4	10.25	30.76	15.79	36.85
dahatonda	13	4	0	0	23.07	69.23
banpimpla	43	7	0	2.32	12.5	42.5
kharbi	31	7	45.16	70.96	64.29	78.57

Fig. C6 A summary of output map 2 compared with input map 0 for chosen villages across 5 Talukas

which were responsible, thereby helping in their resolution. Overall, the pilot has demonstrated that significant improvements in the accuracy and effectiveness of land records may be achieved by the automatic generation of base maps. These reduce the case load substantially and provide a foundation for more detailed reconciliation and land transfers in the future. Based on these results, the department has initiated a long-term collaboration to apply these algorithms and integrate them into the overall modernization process for the state.

In conclusion, leveraging Google farm plots and advanced GIS algorithms presents a promising avenue for modernizing land records and enhancing land management practices. By addressing challenges in data reconciliation, ownership-possession harmonization, and accurate representation of land use, this research contributes to the advancement of digital solutions for sustainable land governance.

C.2 Crop Mapping in Telangana

C.2.1 Introduction

The collaborative approach of Google's ALU research has been pivotal in providing a roadmap for addressing the unique challenges of India's agricultural landscape. The Government of Telangana along with local startup TeamUp, evaluated ALU algorithm's outputs. Field visits and data collection exercises were conducted in 19 villages in the Karimnagar district, to gather feedback from local farmers and stakeholders. This ground-level validation has been crucial in assessing the accuracy of field boundary mapping, identifying operational challenges in implementation & deployment, gathering consultative feedback from farmers and other stakeholders, and areas for improvement.

C.2.2 Field Validation Process and Results

TeamUp, under the guidance of Department of Agriculture, Government of Telangana, undertook the task of integrating the Google ALU algorithm's outputs into their existing crop booking application, which is used for manual monitoring of agricultural fields by agricultural extension workers. During field observations in 19 pilot villages, challenges such as under-segmentation, over-segmentation, and boundary errors were identified. These errors were carefully studied and reported as feedback, highlighting areas for further improvement in the ALU Algorithm's performance.

We observed that 2.19% fields were under-segmented, 7.45% fields were over-segmented for our use case. 8.18% fields were reported to have boundary errors (due to either temporal shift in land use, or lack of clarity of instructions in reporting process).

C.2.3 Future Work

Future iterations of the algorithm should aim to minimize segmentation and boundary errors, providing more accurate field boundary mapping. This will enable more precise and actionable insights for sustainable agricultural practices. The process developed for field validation, integrated with the application, must be scaled for continuous improvement. This methodology can serve as the foundation for both validating and deploying the technology on a larger scale.