Evaluation of computer vision pipeline for farm-level analytics: A case study in Sugarcane

SAMBAL SHIKHAR, Plaksha University, Mohali
RAJIV RANJAN, Plaksha University, Mohali
AMAN SA, Plaksha University, Mohali
ANSHIKA SRIVASTAVA, Plaksha University, Mohali
YASH SRIVASTAVA, Plaksha University, Mohali
DINESH KUMAR, Plaksha University, Mohali
SHASHANK TAMASKAR, Plaksha University, Mohali
ANUPAM SOBTI, Plaksha University, Mohali

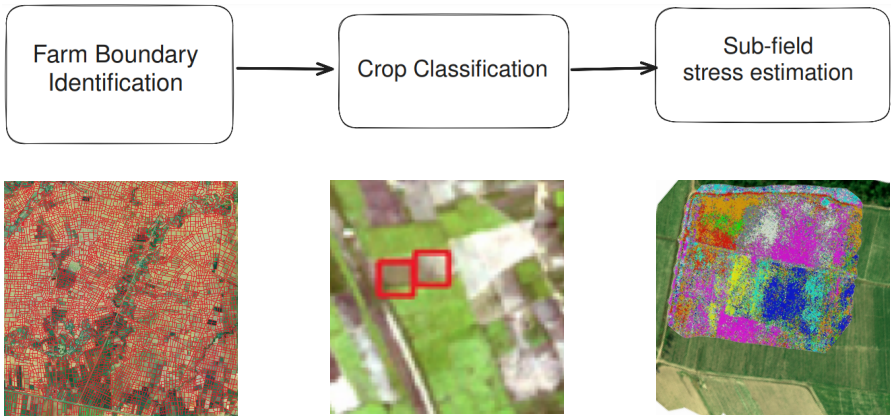


Fig. 1. From left to right: An entire village segmented into farm boundaries, image of one farm from Sentinel2 (10m) satelite (for crop classification), stress estimation in the field on high resolution drone imagery. For a scalable computer vision based pipeline where various stakeholders (governments, financial institutions, farmers) can be served, algorithms for farm boundary identification and crop classification are created on low resolution imagery. Sub-field stress estimation is studied using both high resolution satellite imagery and drone data. We evaluate (and improve) the state of the art in each of these system components, specifically for the sugarcane crop.

Analyzing agricultural imagery for farm level insights has been an active area of research in the recent times. For providing the necessary information to stakeholders - be it farmers, financial institutions or governments, various computer vision tasks have to come together. For example, to provide information to a farmer about crop stress in their farm, accurate localization of the farm, identification of the crop type and a monitoring of the field's micro-climate must be done together. In this work, we set performance benchmarks for three computer vision tasks - farm boundary detection, crop classification and sub-field stress estimation with different

Manuscript submitted to ACM

modalities of images - Sentinel2, PlanetScope and Drone Imagery. We use public dataset benchmarks for farm boundaries and crop classification and do a controlled field study on a large sugarcane farm in Uttar Pradesh, India for the stress estimation.

Our work benchmarks farm boundary detection for small farms with state of the art deep learning algorithms achieving a dice score of 67%, improves the sugarcane classification accuracy by 10% coming to 98% and demonstrates an accuracy of 72% for water and nitrogen stress estimation.

CCS Concepts: • **Computing methodologies** → *Machine learning*;

Additional Key Words and Phrases: Sugarcane, Crop classification, Farm Boundaries, Stress Estimation

ACM Reference Format:
Sambal Shikhar, Rajiv Ranjan, Aman Sa, Anshika Srivastava, Yash Srivastava, Dinesh Kumar, Shashank Tamaskar, and Anupam Sobti. 2024. Evaluation of computer vision pipeline for farm-level analytics: A case study in Sugarcane. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS '24), July 8–11, 2024, New Delhi, India.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3674829.3675079

1 INTRODUCTION

Computer vision has been a promising method to advance information and communication technologies (ICT) for the benefit of various stakeholders in agriculture - the farmer, the financial institutions (loan/insurance) and governments (land records, regulation, policies etc.) [8, 25]. Satellites today are available at multiple resolutions, both temporal and spatial, thus offering a range of capabilities for such ICT technologies. Some examples are Sentinel-2[15] available every week at 10m resolution, PlanetScope[22] available almost every day at 3m resolution and Maxar[1] available on-demand through the satellite themselves and almost a half yearly update through the Google API[2]. However, a practical tool for any of these stakeholders requires that all components of the system perform at a reasonable level of accuracy. It is also important to understand which type of images (modalities) are required for which kind of tasks. In this work, we study a pipeline for farm level insights from satellite and drone images.

Government institutions/departments require data both at an aggregate level - e.g., the area growing a certain crop and at a personalized level, e.g., to check land records. Both these use cases require geolocated boundaries of a field and a classification of the type of crop. Parcel Detection is the task of identifying the boundaries of a crop that is growing in a contiguous space together. Percello *et. al.*[23] have shown the usage of 10-m resolution Sentinel 2 images for parcel detection. To test for geographical generalization, Wang *et. al.*[33] provide a 10000 boundaries dataset spread throughout India. With parcel detection, applications such as land records which require accurate area measurements can be enabled. Downstream applications such as crop classification also require accurate field boundaries.

Financial institutions such as banks and insurance companies provide loans or insurance to clients. However, in case of farmers, the cost of doing a physical survey to establish land records/monitor crop health is too high in comparison to the amounts being disbursed thus making the process infeasible. Satellite imagery can enable both up-to-date information about the field as well as information on historical performance of the land/farmer. This can enable small scale financial applications by providing the information to the financial institution at a fraction of the cost.

For the farmers themselves, a system which understands the agro-ecological zone for a farm as well as classifies which crops are being sown can provide a personalized package of practice to farmers. Such advisories can reinforce best practices for a crop and provide information about the specific weather and soil conditions for the particular crop.

In our work, we explore how each of these applications can be enabled and through which modalities, i.e., through which satellite/drone imagery. We obtained both image data as well as agronomic data by setting up a 16 acre farm where differential treatment of the field was done to simulate water/nitrogen stresses in the test blocks and build models

to detect the same. We establish strong benchmarks on public/private datasets through state of the art algorithms to provide an accurate picture of the current system accuracies which could enable downstream applications for farmers, financial institutions and governments. We compare our algorithms to state of the art in the tasks which have been evaluated before and show improvements to prior art. We benchmark multiple agriculture parcel segmentation algorithms on AI4small boundary dataset [23] , for crop classification using Sentinel2 images and 72% accuracy for water and nitrogen stress images using drone data collected at 2cm resolution.

2 RELATED WORK

In this section, we enumerate different datasets available and the state of the art methods used for the system components of parcel detection, crop classification and sub-field stress identification. Section 3 covers the specific formulations and data used in our experiments.

Farm Boundary and Parcel Detection. Farm boundary and parcel detection are related tasks defined in slightly different ways. In the farm boundary task, an explicit farm boundary mask is used as the ground truth which marks the area separating different fields. On the other hand, the parcel segmentation task uses the farmland directly as a unit of farming, i.e., the same crop growing together without any boundaries in between is a single parcel. AI4Boundaries[13] is a large farm dataset from Europe, AI4SmallFarms[23] provides annotations of smaller farms from Cambodia and Vietnam and Wang *et. al.*[33] provide sparse yet geographically distributed 10000 samples of farm boundaries across India. Aung *et. al.*[4] annotated 2000 tiles across France and collected 3 month sentinel composites for each tile.

Methods for farm boundary detection typically use a semantic segmentation approach where each pixel of the input image is classified into a binary class - boundary/non-boundary. The evolution of semantic segmentation in computer vision and remote sensing has been significantly influenced by the development of convolutional neural network architectures, with UNet introducing the encoder-decoder paradigm for detailed image annotation through multi-level feature representation and dense classification. Taravat *et. al.*[26] improved UNets with skip connections addressed the challenges of vanishing or exploding gradients, enhancing training efficiency for deeper networks. FracTAL ResUNet[31] enhances semantic segmentation by combining ResUnet architecture with FracTAL Attention layers. Long *et. al.*[19] introduce a multi-task neural network to improve the boundary delineation task by integrating a single encoder with three parallel decoders for field identification, boundary prediction, and distance estimation. SEANet[18] is a semantic edge-aware multi-task neural network that enhances the delineation of agricultural parcels from remote sensing images by integrating mask prediction, edge prediction, and distance map estimation, achieving superior accuracy and transferability across various datasets and resolutions. We benchmark some of these state of the art networks on the AI4SmallFarms dataset[23].

Crop Classification. For the crop classification task, both field boundaries (along with images) and the corresponding class labels are typically provided. Some examples include HLS Multitemporal Crop Dataset[7] using a combination of Sentinel-2 and Landsat images in a harmonized format, AgriFieldNet India [37] using Sentinel-2 chips for crops of 13 classes in India while annotating pixel wise crop class, and a UAV dataset from Rwanda collected by Chew *et. al.*[11] for 19 different crop cover types in 5 distinct agro-ecological zones. The CropHarvest [27] dataset offers 90,480 geo-referenced crop data points, integrating satellite time series with binary crop/non-crop and crop type labels. TimeSen2Crop [35] is a pixel based dataset made up of more than 1 million samples of Sentinel 2 Time Series associated with 16 crop types.

Existing crop classification methods span from supervised traditional models like SVM, RF, and MLP, which leverage spectral data and vegetation indices without exploiting time series relationships, to advanced supervised deep learning techniques such as RNNs and CNNs, adept at processing sequential and spatial data. Traditional methods [16, 36], while less complex and quicker, rely heavily on spectral features and statistical analyses of vegetation indices, often employing curve fitting techniques to model crop growth characteristics for improved classification accuracy. On the other hand, deep learning approaches, particularly RNN and LSTM, excel in capturing temporal dynamics but face challenges in parallel computation and large-scale mapping. CNNs offer computational efficiency through sparse interactions and parameter sharing Virnodkar *et. al.*[30] Kussul *et. al.*[17]. Hybrid models[20, 34] combining RNN and CNN attempt to harness both temporal and spatial feature extraction but are limited by their focus on local features, neglecting the comprehensive representation of crop growth stages. Transformers have also been explored to better model spatio-temporal relationships in the crop cycle[32].

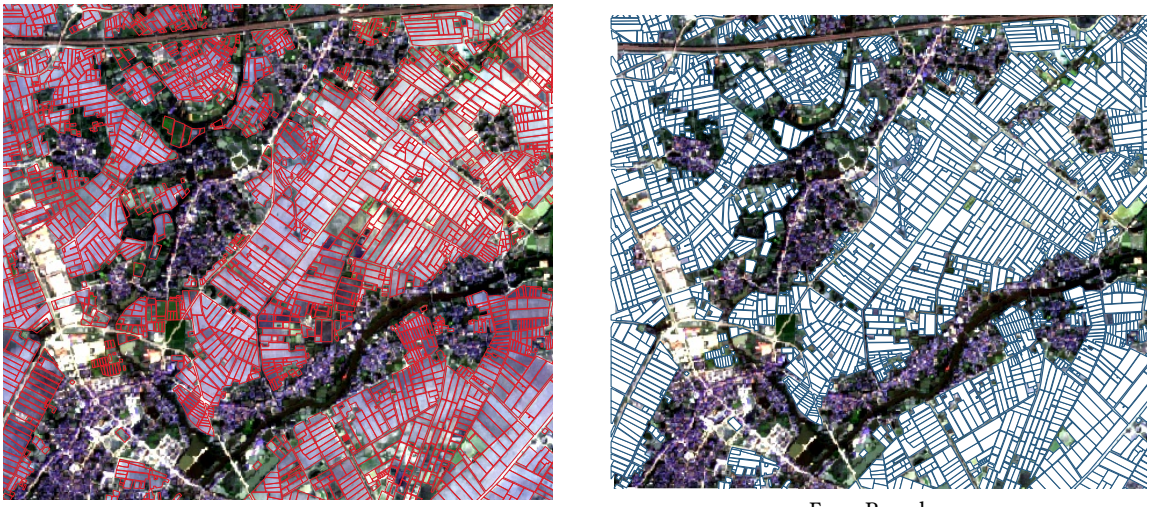
Crop Stress Detection. Crop stress detection uses various methods, e.g., physics based approaches to estimate evapotranspiration through water balance approaches[3]. Behmann *et. al.*[6] introduce a method combining unsupervised and supervised learning with hyperspectral imaging to detect early stages of drought stress in barley by identifying stress-related changes and progressive senescence at the pixel level, significantly outperforming traditional NDVI methods. There have been efforts to classify stress using images from satellite using deep learning approaches [14, 28]. Dadsetan *et. al.*[12] leverages high-resolution aerial imagery and deep learning models, particularly a U-Net combined with a convolutional LSTM, to detect and predict Nutrient Deficiency Stress (NDS) in crops.

Pipelines for stakeholders. While there hasn't been a lot of work on evaluation of pipelines specific to applications for governments, farmers or insurance companies, some concept demonstrators have been done for different stakeholders. For example, in Lattice[9], the authors evaluate a multimodal approach where farm level data is collected with IoT sensors, and cameras and merged with satellite based information. However, the focus is on IoT based information gathering. This type of solution is applicable mostly in large farms with high incomes. Our focus is on a low cost scalable solution for larger areas, e.g., entire states in the country. Another system, eCrop[21] is designed to find damaged areas in a crop. The system in this case also requires IoT deployment and a lot of intervention from the farmer which is not possible for a state wide low resource deployment such as India.

3 TASK DEFINITION, DATASETS AND EVALUATION

3.1 Farm boundary detection

3.1.1 Datasets: Farm boundary. We leverage AI4SmallFarms[23] dataset, developed for assessing the accuracy of crop field boundary delineation in Southeast Asia's smallholder farms. It encompasses 439,001 digitized field polygons across 62 non-overlapping tiles, each approximately 5×5 km, spread throughout Vietnam and Cambodia. The Field boundaries are published in vector format as polygons and polylines, allowing for easy integration with Earth Observation (EO) satellite data. The digitization process included a rigorous multi-step quality control approach to ensure the highest data accuracy, involving visual interpretation of Sentinel-2 (S2) and very high-resolution Google Map (GM) RGB images. The resulting dataset has 62 non-overlapping sentinel-2 tiles with 43 tiles in the training set and the rest divided into valid and test.



Farm Boundaries Farm Parcels

Fig. 2. Example images showing farm boundary and parcel detection tasks. Parcels are focused on learning green uniform areas while boundaries learn the gaps in the green areas.

3.1.2 Evaluation. The Dice score and Intersection over Union (IoU) are two commonly used metrics to evaluate the performance of image segmentation tasks. Both metrics provide a quantitative measure of how well the predicted segmentation aligns with the ground truth.

Intersection over Union (IoU) IoU, also known as the Jaccard Index, measures the overlap between two boundaries. Farm boundary segmentation quantifies the overlap between the predicted boundary and the ground truth. Assuming G is the ground truth parcel and P is the predicted parcel, the IOU is defined as:

$$\text{IoU} = \frac{P \cap G}{P \cup G} \tag{1}$$

where $P \cap G$ represent the area common between the predicted and ground truth boundary and $P \cup G$ represents the area in either P or G boundary.

The Dice score, or Dice coefficient, is similar to IoU but tends to give higher values for partial overlaps, making it a more optimistic measure of similarity. It is defined as

$$\text{Dice} = \frac{2 \times P \cap G}{P \cup G} \tag{2}$$

3.2 Crop Classification

3.2.1 Datasets: Crop Classification. The CaneSat dataset addresses the binary classification between sugarcane and non-sugarcane crops in agriculture using cloud-free remote sensing (RS) data from the ESA's Copernicus program, with preprocessed Sentinel-2 images to remove atmospheric, radiometric, and geometric anomalies. The dataset represents four talukas to capture the variance in sugarcane growth due to factors like irrigation, soil properties, and climate. It includes two classes—sugarcane covering all growth phases, and non-sugarcane comprising various land covers. Image patches are 10x10 pixels, derived from Sentinel-2 imagery at 10m resolution, and ground truth data is collected via field surveys and GPS devices, ensuring precise geometric representation. The dataset, with 1627 samples covering

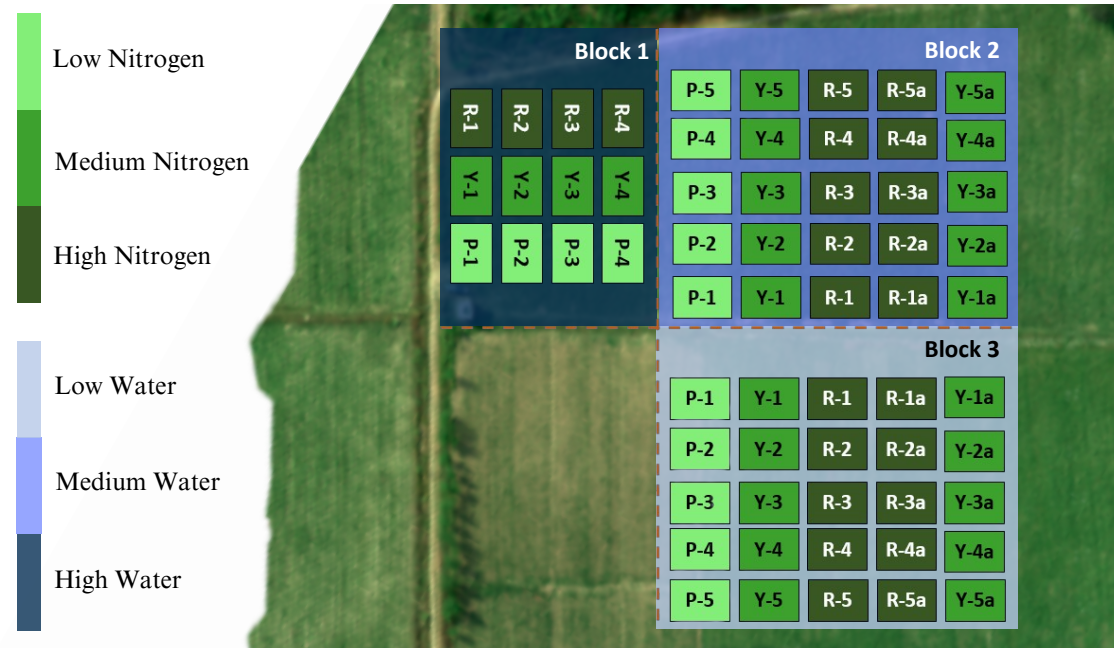


Fig. 3. The test farm is divided into three zones by the amount of water application level (Low, Medium, High Water) - indicated through shades of blue background. Within each zone, there are blocks for low, medium and high nitrogen levels - indicated through shades of green.

16.27 hectares, distinguishes between 87,000 sugarcane and 75,700 non-sugarcane pixels, provided in both jpg and georeferenced tif formats with six spectral bands.

3.2.2 Evaluation. We evaluate the canesat test data accuracy on multiple CNN architectures and custom feature engineered Random forest model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{3}$$

where, TP, TN, FP, and FN refer to true positives, true negatives, false positives and false negatives respectively. The positive class is assumed to be sugarcane. Therefore, TP represents a sugarcane pixel being detected as sugarcane. The dataset and evaluation are at the pixel level. Therefore, the accuracy represents the ratio of total correctly labelled pixels over the total pixels.

3.3 Stress Detection

3.3.1 Farm Setup and Agronomic Data. We set up a 16-acre test farm (see Figure 3) in Lakhimpur District, Uttar Pradesh, India. The area falls in the Upper Gangetic Plains agro-climatic zone in India. The farm was set with two objectives - water and nitrogen stress assessment. For nitrogen stress estimation, the *P*, *Y* and *R* blocks were supplied with 0.25kg, 0.5kg and 1kg of fertilizer for block 1 and 0.2kg, 0.4kg and 0.8kg of fertilizer for blocks 2 and 3. The water levels were 50%, 100% and 150% of the typical application for Low, Medium and High water blocks. These quantities were decided based on the recommendation of an agronomist. On this test farm, we did monthly drone flights and collected drone images at 2cm resolution (37.5m drone height). We use this drone dataset for the image based analysis.

Fig. 4. Leaf samples collected from the farm are stored in a box to preserve moisture (left). Within 24 hours, leaf samples are dried in an oven (right). Weights are taken at both stages to estimate leaf water content.

In this study, two kinds of agronomic data for sugarcane were collected - Leaf water content and cane height.

Leaf water content (LWC) is computed using the formula given below, where "Wet Leaf" is the weight of the leaf when wet (just after cutting leaves from canes), and "Dry Leaf" is the weight of the leaf when it is dry (dried using a heater).

$$\text{LWC} = \frac{\text{Wet Leaf Weight - Dry Leaf Weight}}{\text{Dry Leaf Weight}} \tag{4}$$

Fresh cane's leaves were randomly collected from each block. Within 24 hrs, it was weighed and put into an oven for 72 hrs at 60°C temperature for drying as shown in Figure 4. After drying it was weighed again.

Water & Nitrogen Level	**Abbreviation**
Low Water, Low Nitrogen	LW_LN
Low Water, Medium Nitrogen	LW_MN
Low Water, High Nitrogen	LW_HN
Medium Water, Low Nitrogen	MW_LN
Medium Water, Medium Nitrogen	MW_MN
Medium Water, High Nitrogen	MW_HN
High Water, Low Nitrogen	HW_LN
High Water, Medium Nitrogen	HW_MN
High Water, High Nitrogen	HW_HN

Table 1. Combinations of Water and Nitrogen Levels

In Table 2, LWC for three blocks has been analysed i.e., B1, B2 and B3 which are high water, medium water and low water blocks respectively. The LWC for three dates (till 29-06-2023) supports that the differential water application

Blocks	01/05	29/05	29/06	17/08
B1	1.898	2.427	3.085	5.475
B2	2.053	2.390	2.666	5.740
B3	2.033	2.267	2.558	5.742
Rainfall (mm)	7.05	36.09	71.66	276.82

Table 2. The average leaf water content recorded shows that the water stress induction was successful. The crop recovered from water stress after heavy rains in July. The dates are in dd/mm format.

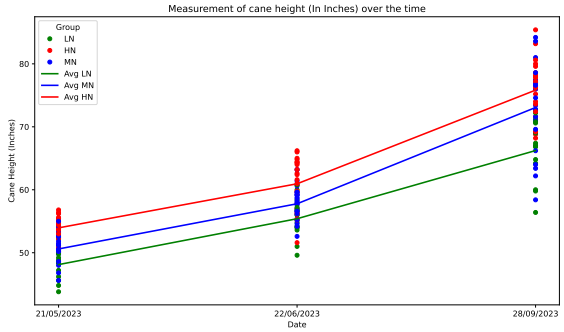


Fig. 5. The cane height was consistently higher for higher nitrogen inputs.

resulted in increasing water stress for B1, B2 and B3. The data shows a median across 5 samples from each block. High rainfall in July and August reverses the order of leaf water content after the third date (29-06-2023).

Cane Height is the height of the cane without the leaves. The measurement is taken as the average for five random canes collected from each block (as shown in Figure 3). In Figure 5, the increment in cane's height can be easily seen. The average cane height for the high nitrogen (HN) block in the red line lies always at the top, for the medium nitrogen (MN) block in blue at the middle and the low nitrogen (LN) block in green at the bottom of the graph. As green, blue and red are the low, medium and high fertilizer blocks respectively. The cane height graph shows the impact of the use of levels of fertilizer.

Yield At the time of harvest, the cane harvested in each crop was weighed with an industrial grade weighing machine. We use yield as a mild indicator of stress, i.e., stressed regions are expected to have lower yields. The production is analysed across all blocks. In Figure 6 we can see that increasing nitrogen application linearly increases the yield across all water levels. The high water (medium and high nitrogen) blocks were exposed to some animals which explains the unexpected loss in yield. For a constant nitrogen level, the water stress developed in the crop pre-monsoon results in much lower yields. For example, the LW_LN, MW_LN and HW_LN have a yield of 158, 182 and 193 kg/acre.

3.3.2 Datasets: Stress Detection. As discussed in Section 3.3.1, we utilize a high-resolution drone image dataset representing various agricultural land attributes taken from a 35m distance. Each block within the dataset is characterized by specific conditions, such as 'High Water and Low Nitrogen,' among 9 total categories made using all possible combinations of 'high,' 'medium,' and 'low' in the case of Nitrogen and Water each (refer Figure 3). To facilitate the training of our convolutional neural networks (CNNs), we further divide (Figure 7) each block into smaller, more manageable 'chips' of 32×32 pixels corresponding to $64cm \times 64cm$ on the ground. When generating the dataset, the

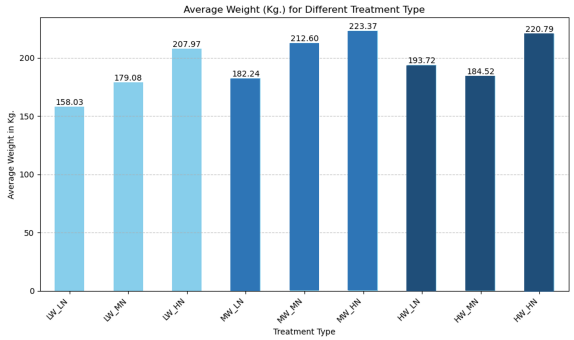


Fig. 6. Yield comparison across different blocks of the test farm. Nitrogen level has a direct correlation with yield. With the same nitrogen level, higher water applications results in better yield even though the water stress lasted pre-monsoon season only.

chips being generated have a certain amount of overlap between them. This granular division allows the CNN to learn from a more extensive range of localized patterns. Each chip serves as an independent data point, labelled with the attributes of its corresponding block as shown in Figure 7. The ratio of the overlapping area is referred to using the word 'stride.' The data taken from the drone has 5 spectral bands, namely red, blue, green, NIR(Near Infrared), and RE(Red Edge). We then extracted the Green Leaf Index (GLI) and Normalized Difference Vegetation Index (NDVI) [29] corresponding to each pixel from the five bands as given in the following formulas.

$$\text{GLI} = \frac{2 \times \text{green} - \text{red} - \text{blue}}{2 \times \text{green} + \text{red} + \text{blue}} \tag{5}$$

$$\text{NDVI} = \frac{\text{NIR} - \text{red}}{\text{NIR} + \text{red}} \tag{6}$$

3.3.3 Evaluation Metrics. Our evaluation of the stress detection model encompasses several key metrics: precision, recall, accuracy, and F1-score. These metrics collectively offer a comprehensive understanding of the model's performance in detecting combined water and nitrogen stress, as well as its effectiveness in the individual water and nitrogen stress classifiers.

(1) **Accuracy:**
Accuracy is the total number of true positives across classes in proportion to the total number of points.

$$Acc = \frac{\sum_{i=1}^{M} TP_i}{\text{Total Points}} \tag{7}$$

where $i \in [1, M]$ corresponds to the different water/nitrogen categories. For the combined stress model, $M = 9$ while for separate water and nitrogen stress levels, $M = 3$.

(2) **Precision:**
Precision is an indicator of the model's capability to accurately label a certain class. For i^{th} class, high precision (P_i) signifies a low rate of false positives:

$$P_i = \frac{TP_i}{TP_i + FP_i} \tag{8}$$

(3) **Recall:**

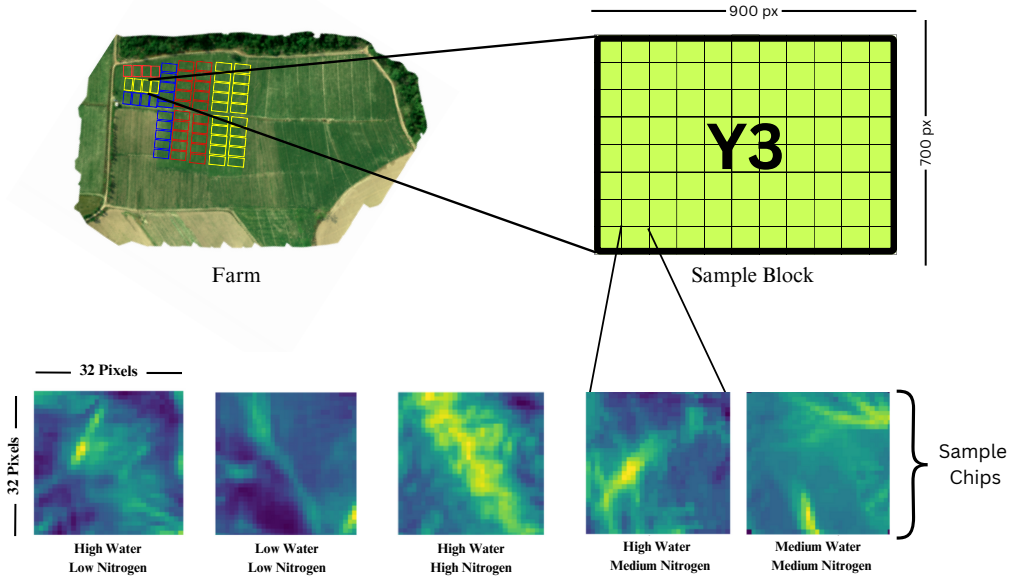


Fig. 7. Data loader for stress detection

Recall measures the model's sensitivity in identifying all instances of a given stress level. High recall (R_i) of a model demonstrates efficiency in capturing the majority of stress occurrences:

$$R_i = \frac{TP_i}{TP_i + FN_i} \tag{9}$$

(4) **F1-Score:**

F1-Score harmonizes precision and recall, providing a balanced evaluation of the model's overall effectiveness. In the combined model, a high F1-Score ($F1_i$) indicates a well-balanced model

$$F1_i = \frac{2 \cdot P_i \cdot R_i}{P_i + R_i} \tag{10}$$

The reported scores are averaged across all M classes.

4 COMPUTER VISION PIPELINE EVALUATION

4.1 Farm Boundary

We benchmark the Ai4SmallFarms[23] dataset parcel area segmentation task using UNet [24] and DeeplabV3 [10] architectures trained on all the 43 training images. These are large images with 100s of parcels per image as shown in Figure 2. Further, we also benchmark Ai4smallfarms using state of the art multitask architectures BsiNet [19] and FAUNet [5]. The encoder backbone for UNet uses pretrained Resnet (with Imagenet) weights. We use the pretrained EfficientNet-b7 weights for DeeplabV3 encoder backbone. We train BSINet and FAUNet from scratch with both boundary edge delineation and parcel area segmentation tasks. As shown in Table 3, the DeepLabV3 architecture performs much better in comparison to UNet because of DeepLabV3's ability to capture multi-scale contextual information using

Architecture	Backbone	IoU	Dice Score
UNet	Resnet-50	0.48	0.641
DeepLabV3	EfficientNet-b7	0.51	0.672
BSINet	-	0.486	0.645
FAUNet	-	0.493	0.651

Table 3. DeepLabV3 with its limited parameters and a multiscale contextual approach performs best on the AI4SmallFarms[23] dataset.

Architecture	Accuracy
ResNet-50	85.65%
DenseNet201	87.70%
RF with Temporal Features	98.2%

Table 4. A random forest classifier with NDVI-based features over a temporal window of an year is able to outperform convolutional architectures which aim to directly classify sugarcane using image at a single timestamp.

architectural features like atrous spatial pyramid pooling. As the backbone is initialized with efficientnet-b7 and the training data is limited, DeepLabV3 benefits from transfer learning and performs better than the multitask architectures which are trained from scratch. The Dice Score of 67.2 is the state of the art score on parcel extraction tasks for small farms[23].

4.2 Crop Classification

Virnodkar *et. al.*[30], the authors of CaneSat, primarily used convolutions based architectures, however, it fails to account for temporal properties exhibited by crops. We created temporal features by extracting sentinel-2 images for each plot corresponding to 12 months of the year as Sugarcane is a full year crop and we want to account the entire crop cycle. NDVI map corresponding to each month for the respective training plot images are then extracted. The images are then filtered based on cloud percentage value and only images with less than 5% of cloud cover are used. After filtering we are left with 7 months of cloud free images and corresponding NDVI maps. The mean values of the corresponding NDVI maps are treated as features for each plot which are fed into a Random Forest classifier to classify a given plot image as Sugarcane/Non-Sugarcane. We found that our simple time series approach is able to get an accuracy of 98.2% as compared to the convolution approach by Virnodkar*et. al.*[30] which yields an accuracy of 87% using a DenseNet architecture. This also demonstrates the effectiveness of vegetation index NDVI to capture the vegetation effectively.

4.3 Crop Stress

As described in Section 3.3.1, the stress analysis was done by setting up a test farm and taking agronomic measurements over the entire crop cycle (a year). In this section, we share results from both satellite based analysis and drone image based analysis. The goal of the analysis is to provide real-time insights to a farmer on any stresses present in the field. Therefore, the analysis is done on a single timestamp (21^{st} June 2023.)

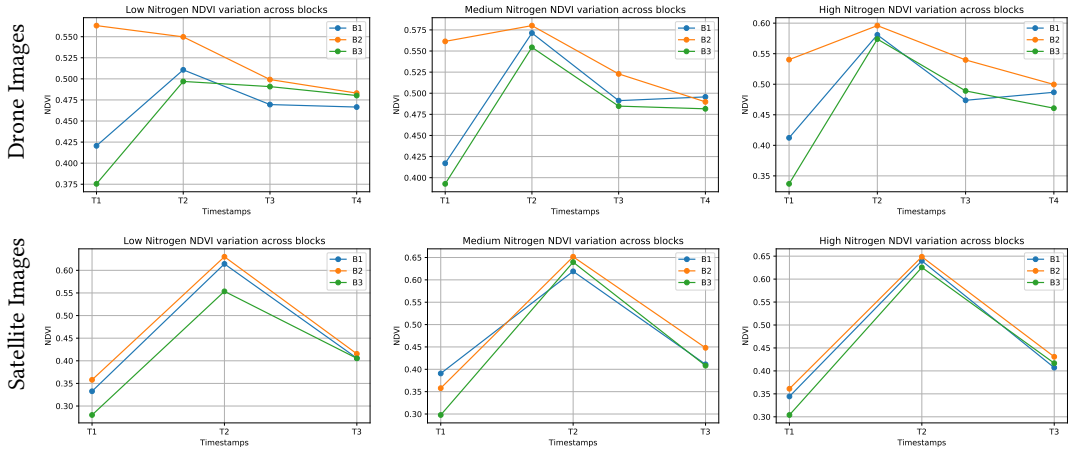


Fig. 8. NDVI initially starts with different values but converges towards the harvest period. T1=Apr, T2=May, T3=June, T4=Aug. Satellite images were not available at T4 from the satellite due to cloud cover.

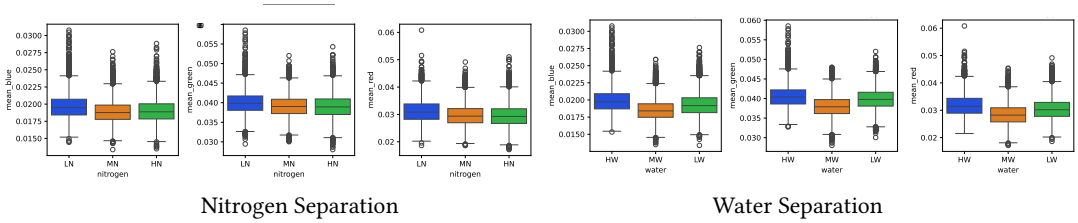


Fig. 9. Water and nitrogen both show overlapping bands in raw data in 32×32 chips. Even though distributions show slightly different medians, the high overlap indicates that raw separation of classes will not be possible.

4.3.1 Initial Analysis. To start with, we validate if the image had a sufficient distinctive signal for the models to learn from. Our initial analysis of the agricultural data using statistical methods revealed distinct patterns in NDVI values across blocks with varying water and nitrogen levels. As shown in Figure 8, the mean NDVI of medium water blocks (B2) was consistently higher over time, whereas high water blocks (B1) showed lower NDVI values. Significant fluctuations in NDVI were observed in B3, indicative of variable growth conditions.

As the season progressed, the variance in NDVI across all blocks reduced, suggesting a homogenization of vegetative states. High and medium nitrogen blocks exhibited better NDVI values, affirming the positive impact of nitrogen. Importantly, the initially large discrepancies in NDVI between different treatments narrowed over time. Band Separability Analysis further validated this clear separability observed between low water and other blocks, as well as across different nitrogen levels in Figure 9.

4.3.2 Satellite based analysis. Planetscope 4-band image has a resolution of 3 meters. Therefore, one test block for each class contains approximately 20 to 25 pixels. We split the 62 blocks from all classes into train and test. We used pixels of 49 blocks for training and 13 blocks for validation. There are a total of nine classes as defined in Figure 3. We used the NDVI values of each block as a feature vector. We used a random forest (RF) classifier to find the appropriate class of each pixel and got an accuracy of 38.46%.

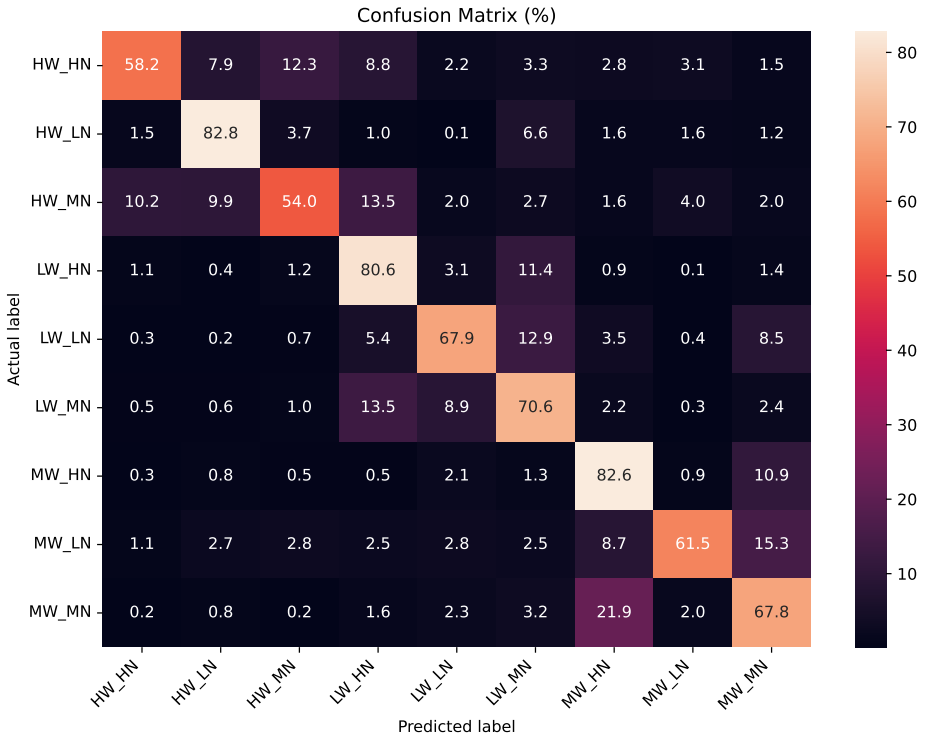


Fig. 10. Confusion matrix for various stress levels

4.3.3 Drone-based analysis. As described in Section 3.3.2, we broke down the drone image captured into 32×32 chips. We further broke it down into smaller 8×8 pieces to generate a random forest classifier. For each of the 8×8 pieces, we took the average NDVI and GLI of each piece. With these 32 features - 16 NDVI values and 16 GLI values, a max depth of 10, min samples split of 5 with 500 estimators gave us a baseline accuracy of 23.0%. We also used Convolutional Neural Networks (CNNs) for our analysis, with a stride of 0.35 and approximately 1.4 lakh training chips we achieved an accuracy of 71.8%. To evaluate the robustness of the model, we incorporated various data augmentation strategies. Table 5 demonstrates the effectiveness of different combinations of flip (horizontal and vertical), rotation (10 degrees), and intensity (gaussian noise, colour jitter and Gaussian blur) augmentations with stride values set at either 0.35 or 0.51 and a probability of augmentation at either 0.3 or 0.6. These adjustments allowed us to generate a more varied dataset for training the CNNs and prevent overfitting. We show the results of these models in Figure 11. As one can observe, the accuracy without the augmentations was 72%. We were able to verify that even with augmentations, the accuracy remains in a similar range indicating that the model was able to learn the classifier with reasonable accuracy without augmentations.

Figure 10 shows a confusion matrix that represents the performance of the best classification model in percentage terms. Each cell in the matrix indicates the percentage of predictions for the class shown by the column label, given the actual class shown by the row label. For instance, the correct prediction rate for 'High water and high nitrogen'

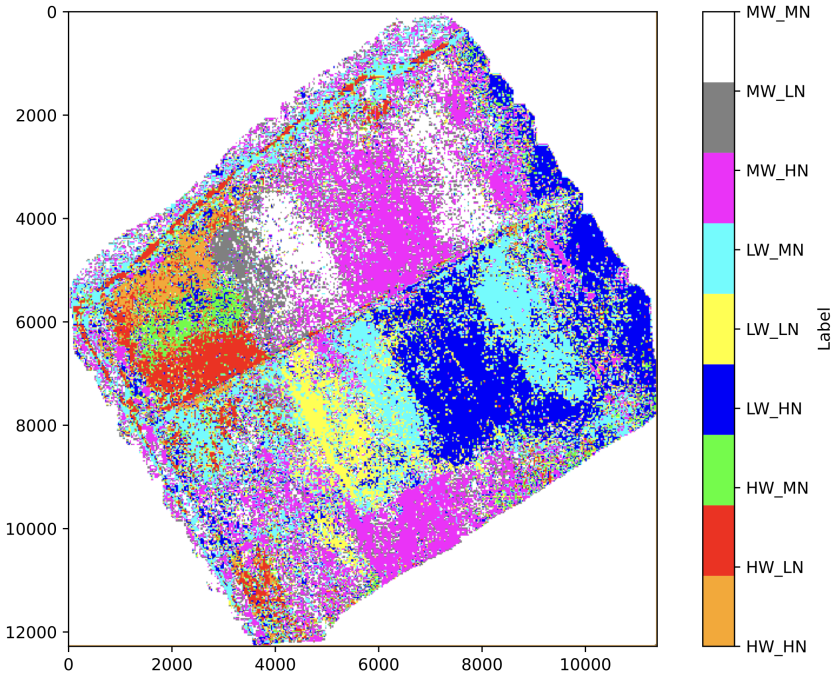


Fig. 11. Results of the classification are overlayed over the drone image collected. The inference is done over a window of 32×32.

(HW_HN) is 58.2%. The classification model works well in identifying 'Low water and low nitrogen' (LW_LN) at 67.9%, implying a solid understanding of this class's features. However, it often mistakes 'Low water and medium nitrogen' (LW_MN) for 'Medium water and medium nitrogen' (MW_MN), indicating an area that needs improvement. The high confusion rates between classes, especially where water levels vary but nitrogen levels stay medium, suggest that the model struggles with fine differences in varying levels of water.

Finally, we evaluate if instead of the 9 categories, it helps the classifier learn better if we use only 6 categories and separate the classifiers for nitrogen and water.

Water Stress Classifier. Our water stress classification model effectively distinguishes between Low and Medium Water stress levels, achieving precision scores of 0.621 and 0.745, respectively. Notably, it exhibits high recall rates, with 0.861 for Low Water stress areas and 0.606 for Medium Water stress, indicating its ability to accurately identify regions experiencing moderate water scarcity. However, the model encounters challenges in the classification of High Water stress regions, as reflected in a lower precision of 0.301 and a recall of 0.268. This difficulty implies limitations in recognizing areas with excess water, possibly due to the complex spectral signatures associated with high water content. The overall accuracy of the model is 0.63, suggesting a reasonable performance.

Augmentation	Stride	Probability	Val Accuracy
None	0.35	0	0.72
None	0.51	0	0.65
Flip and rotation	0.35	0.3	0.70
Flip and rotation	0.35	0.6	0.69
Flip and rotation	0.51	0.3	0.64
Flip and rotation	0.51	0.6	0.65
Flip, rotation and intensity	0.35	0.3	0.69
Flip rotation and intensity	0.35	0.6	0.64
Flip, rotation and intensity	0.51	0.3	0.63
Flip, rotation and intensity	0.51	0.6	0.51

Table 5. We applied various augmentations to our CNN model to ensure that the results were not overfit on the training data. With more data, the validation accuracy reduced slightly but we found results for the vanilla training without augmentations to be comparable to the training with augmentation.

Stress Level	Precision	Recall	F1-Score	Support
Low Water	0.621	0.861	0.722	6511
Medium Water	0.745	0.606	0.668	11847
High Water	0.301	0.268	0.28354	3641
Accuracy				0.63

Table 6. When trained only on water labels, the low and medium water classes show much higher F1 score than high water. Despite the high variance in raw band values (Fig 9), there seem to be features that are able to discriminate the two.

Stress Level	Precision	Recall	F1-Score	Support
Low Nitrogen	0.435	0.262	0.327	4379
Medium Nitrogen	0.526	0.281	0.366	7835
High Nitrogen	0.463	0.772	0.58	8072
Accuracy				0.47

Table 7. When trained on only nitrogen labels, very low F1-scores are observed. Therefore, the joint model is preferred.

Nitrogen Stress Classifier. Examining the nitrogen-based classifier, our findings reveal a varying performance across distinct stress levels. The overall accuracy of the model stands at 0.47, indicating moderate success in classifying nitrogen stress in our dataset. Notably, the model demonstrates effectiveness in detecting regions with High Nitrogen stress, achieving a precision of 0.463 and a recall of 0.772. This suggests its aptitude for identifying areas characterized by elevated nitrogen content. The model's precision and recall for Low Nitrogen stress are 0.435 and 0.262, respectively, indicating limitations in accurately distinguishing areas with lower nitrogen content. Similarly, the classification of Medium Nitrogen stress proves challenging, with a precision of 0.526 and a recall of 0.281. These results demonstrate that the model with separate 9 classes performs better than the separate classifiers. We hypothesize that this is due to lesser intra-class variance within each class for separate classifiers.

5 CONCLUSION

We evaluate a computer vision pipeline for building scalable informatics regarding a farm. We started from farm boundary extraction, and crop classification using public datasets and did deeper experiments on stress classification on a test farm. Here we collected high resolution drone imagery, physiological measurements as well as satellite images of the crop.

Our experiments show that low resolution (10m Sentinel2) images are an efficient modality for farm boundary and classification tasks. It also helps that such imagery is freely available thus reducing the cost of such a solution required to be deployed at scale. We were able to achieve a dice score of 67% with a DeepLabv3 model trained in a supervised setting and a sugarcane classification accuracy of 98.2% using a random forest classifier on temporal NDVI features. We set up a 16 acre test farm and were successfully able to induce water and nitrogen stress in the farm. We were able to take physiological measurements of leaf water content and cane height from the farm to validate these stresses. We then built classifiers for classifying these areas based on drone and satellite images. Analysis of water and nitrogen stress in our test farm indicates that sub-field level information is only captured through high resolution imagery - PlanetScope (to an extent) and more completely through Drone Flights. Temporal models on high resolution imagery such as drones also will be attempted in future work. Another area where models could benefit is calibration mechanisms to incorporate agronomic data based feedback from the ground.

ACKNOWLEDGMENTS

We'd like to extend a big thank you to Mr. Vikram Bhalla for his assistance in securing a test farm within this expansive agricultural landscape. His support has been invaluable to our research.

We also want to express our gratitude to CNH Industrial for collaborating with our university. Their partnership has provided us with valuable resources, financial support and insights, enhancing our academic pursuits in agriculture.

REFERENCES

[1] [n. d.]. Maxar Imagery. https://maxar.com/maxar-intelligence/products/optical-imagery. [Accessed 06-03-2024].

[2] [n. d.]. Overview | Maps Static API | Google for Developers — developers.google.com. https://developers.google.com/maps/documentation/maps-static/overview. [Accessed 06-03-2024].

[3] Mohammad Alavi, Mohammad Albaji, Mona Golabi, Abd Ali Naseri, and Saeid Homayouni. 2024. Estimation of sugarcane evapotranspiration from remote sensing and limited meteorological variables using machine learning models. *Journal of Hydrology* 629 (2024), 130605.

[4] Han Lin Aung, Burak Uzkent, Marshall Burke, David Lobell, and Stefano Ermon. 2020. Farm parcel delineation using spatio-temporal convolutional networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 76–77.

[5] Bahaa Awad and Isin Erer. 2023. FAUNet: Frequency Attention U-Net for Parcel Boundary Delineation in Satellite Images. *Remote Sensing* 15, 21 (2023), 5123.

[6] Jan Behmann, Jörg Steinrücken, and Lutz Plümer. 2014. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* 93 (2014), 98–111.

[7] Michael Cecil, Fatemeh Kordi, Hanxi (Steve) Li, Sam Khallaghi, and Hamed Alemohammad. 2023. HLS Multi Temporal Crop Classification. https://doi.org/10.57967/hf/0955

[8] Ranveer Chandra and Stewart Collis. 2021. Digital agriculture for small-scale producers: challenges and opportunities. *Commun. ACM* 64, 12 (2021), 75–84.

[9] Somali Chaterji, Nathan DeLay, John Evans, Nathan Mosier, Bernard Engel, Dennis Buckmaster, Michael R Ladisch, and Ranveer Chandra. 2021. Lattice: A vision for machine learning, data engineering, and policy considerations for digital agriculture at scale. *IEEE Open Journal of the Computer Society* 2 (2021), 227–240.

[10] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. 2017. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* (2017).

[11] Robert Chew, Jay Rineer, Robert Beach, Maggie O'Neil, Noel Ujeneza, Daniel Lapidus, Thomas Miano, Meghan Hegarty-Craver, Jason Polly, and Dorota S Temple. 2020. Deep neural networks and transfer learning for food crop identification in UAV images. *Drones* 4, 1 (2020), 7.

[12] Saba Dadsetan, Gisele Rose, Naira Hovakimyan, and Jennifer Hobbs. 2021. Detection and prediction of nutrient deficiency stress using longitudinal aerial imagery. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 14729–14738.

[13] Raphaël d'Andrimont, Martin Claverie, Pieter Kempeneers, Davide Muraro, Momchil Yordanov, Devis Peressutti, Matej Batič, and François Waldner. 2023. AI4Boundaries: an open AI-ready dataset to map field boundaries with Sentinel-2 and aerial photography. *Earth System Science Data* 15, 1 (2023), 317–329.

[14] Leonardo Leite de Melo, Verônica Gaspar Martins Leite de Melo, Patrícia Angélica Alves Marques, Jose Antonio Frizzone, Rubens Duarte Coelho, Roseli Aparecida Francelin Romero, and Timóteo Herculino da Silva Barros. 2022. Deep learning for identification of water deficits in sugarcane based on thermal images. *Agricultural Water Management* 272 (2022), 107820.

[15] Matthias Drusch, Umberto Del Bello, Sébastien Carlier, Olivier Colin, Veronica Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment* 120 (2012), 25–36.

[16] Pall Oskar Gislason, Jon Atli Benediktsson, and Johannes R Sveinsson. 2006. Random forests for land cover classification. *Pattern recognition letters* 27, 4 (2006), 294–300.

[17] Nataliia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov. 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters* 14, 5 (2017), 778–782.

[18] Mengmeng Li, Jiang Long, Alfred Stein, and Xiaoqin Wang. 2023. Using a semantic edge-aware multi-task neural network to delineate agricultural parcels from remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing* 200 (2023), 24–40.

[19] Jiang Long, Mengmeng Li, Xiaoqin Wang, and Alfred Stein. 2022. Delineation of agricultural fields using multi-task BsiNet from high-resolution satellite images. *International Journal of Applied Earth Observation and Geoinformation* 112 (2022), 102871.

[20] Chang Luo, Shiyao Meng, Xin Hu, Xinyu Wang, and Yanfei Zhong. 2020. Cropnet: Deep Spatial-Temporal-Spectral Feature Learning Network for Crop Classification from Time-Series Multi-Spectral Images. In *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. 4187–4190. https://doi.org/10.1109/IGARSS39084.2020.9324097

[21] Alakananda Mitra, Anshuman Singhal, Saraju P Mohanty, Elias Kougianos, and Chittaranjan Ray. 2022. eCrop: a novel framework for automatic crop damage estimation in smart agriculture. *SN Computer Science* 3, 4 (2022), 319.

[22] Planet Labs PBC. 2018–. Planet Application Program Interface: In Space for Life on Earth. https://api.planet.com

[23] Claudio Persello, Jeroen Grift, Xinyan Fan, Claudia Paris, Ronny Hänsch, Mila Koeva, and Andy Nelson. 2023. AI4SmallFarms: A Data Set for Crop Field Delineation in Southeast Asian Smallholder Farms. *IEEE Geoscience and Remote Sensing Letters* (2023).

[24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 234–241.

[25] BK Singh, Dulal Chakraborty, Naveen Kalra, and Jaya Singh. 2019. *A tool for climate smart crop insurance: Combining farmers' pictures with dynamic crop modelling for accurate yield estimation prior to harvest.* Intl Food Policy Res Inst.

[26] Alireza Taravat, Matthias P Wagner, Rogerio Bonifacio, and David Petit. 2021. Advanced fully convolutional networks for agricultural field boundary detection. *Remote sensing* 13, 4 (2021), 722.

[27] Gabriel Tseng, Ivan Zvonkov, Catherine Lilian Nakalembe, and Hannah Kerner. 2021. Cropharvest: A global dataset for crop-type classification. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*.

[28] Shadman Veysi, Abd Ali Naseri, Saeid Hamzeh, and Harm Bartholomeus. 2017. A satellite based crop water stress index for irrigation scheduling in sugarcane fields. *Agricultural water management* 189 (2017), 70–86.

[29] Andrés Viña, Anatoly A Gitelson, Anthony L Nguy-Robertson, and Yi Peng. 2011. Comparison of different vegetation indices for the remote assessment of green leaf area index of crops. *Remote sensing of environment* 115, 12 (2011), 3468–3478.

[30] Shyamal S Virnodkar, Vinod K Pachghare, VC Patil, and Sunil Kumar Jha. 2022. CaneSat dataset to leverage convolutional neural networks for sugarcane classification from Sentinel-2. *Journal of King Saud University-Computer and Information Sciences* 34, 6 (2022), 3343–3355.

[31] 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.

[32] Hengbin Wang, Wanqiu Chang, Yu Yao, Zhiying Yao, Yuanyuan Zhao, Shaoming Li, Zhe Liu, and Xiaodong Zhang. 2023. Cropformer: A new generalized deep learning classification approach for multi-scenario crop classification. *Frontiers in plant science* 14 (2023), 1130659.

[33] 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.

[34] Giulio Weikmann, Claudia Paris, and Lorenzo Bruzzone. 2021. Multi-year crop type mapping using pre-trained deep long-short term memory and Sentinel 2 image time series. In *Image and Signal Processing for Remote Sensing XXVII*, Vol. 11862. SPIE, 171–181.

[35] Giulio Weikmann, Claudia Paris, and Lorenzo Bruzzone. 2021. Timesen2crop: A million labeled samples dataset of sentinel 2 image time series for crop-type classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), 4699–4708.

[36] Xiaochun Zhang, Qinxue Xiong, Liping Di, Junmei Tang, Jin Yang, Huayi Wu, Yan Qin, Rongrui Su, and Wei Zhou. 2018. Phenological metrics-based crop classification using HJ-1 CCD images and Landsat 8 imagery. *International Journal of Digital Earth* 11, 12 (2018), 1219–1240.

[37] Zindi. [n. d.]. Zindi — zindi.africa. https://zindi.africa/competitions/agrifieldnet-india-challenge. [Accessed 28-02-2024].