

Rapid Response Crop Maps in Data Sparse Regions

Hannah Kerner*

hkerner@umd.edu

University of Maryland, College Park

Gabriel Tseng*

gabrieltseg95@gmail.com

NASA Harvest

Inbal Becker-Reshef

ireshef@umd.edu

University of Maryland, College Park

Catherine Nakalembe

cnakalem@umd.edu

University of Maryland, College Park

Brian Barker

bbarker1@umd.edu

University of Maryland, College Park

Blake Munshell

bmunshel@terpmail.umd.edu

University of Maryland, College Park

Madhava Paliyam

mpaliyam@terpmail.umd.edu

University of Maryland, College Park

Mehdi Hosseini

mhoseini@umd.edu

University of Maryland, College Park

ABSTRACT

Spatial information on cropland distribution, often called cropland or crop maps, are critical inputs for a wide range of agriculture and food security analyses and decisions. However, high-resolution cropland maps are not readily available for most countries, especially in regions dominated by smallholder farming (e.g., sub-Saharan Africa). These maps are especially critical in times of crisis when decision makers need to rapidly design and enact agriculture-related policies and mitigation strategies, including providing humanitarian assistance, dispersing targeted aid, or boosting productivity for farmers. A major challenge for developing crop maps is that many regions do not have readily accessible ground truth data on croplands necessary for training and validating predictive models, and field campaigns are not feasible for collecting labels for rapid response. We present a method for rapid mapping of croplands in regions where little to no ground data is available. We present results for this method in Togo, where we delivered a high-resolution (10 m) cropland map in under 10 days to facilitate rapid response to the COVID-19 pandemic by the Togolese government. This demonstrated a successful transition of machine learning applications research to operational rapid response in a real humanitarian crisis. All maps, data, and code are publicly available to enable future research and operational systems in data-sparse regions.

CCS CONCEPTS

- Applied computing → Agriculture; • Computing methodologies → Neural networks.

KEYWORDS

agriculture, food security, crop classification, neural networks, Earth observation

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

KDD '20 Humanitarian Mapping Workshop, August 24, 2020, San Diego, CA

© 2020 Association for Computing Machinery.

ACM Reference Format:

Hannah Kerner, Gabriel Tseng, Inbal Becker-Reshef, Catherine Nakalembe, Brian Barker, Blake Munshell, Madhava Paliyam, and Mehdi Hosseini. 2020. Rapid Response Crop Maps in Data Sparse Regions. In *KDD '20: ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) Humanitarian Mapping Workshop*. ACM, New York, NY, USA, 7 pages.

1 INTRODUCTION

Satellite Earth observation (EO) data can provide critical, relevant, and timely information in support of agricultural monitoring and food security [2, 15]. EO can enable crop yield assessment and forecasting [3, 7, 8, 13, 21], inform agricultural commodities markets, enable early warning and mitigation of impending crop shortages, and inform decisions about subsidies in times of crisis such as extreme drought [6] or outbreaks of disease or pests. However, the lack of up-to-date, high-resolution cropland and crop type maps is a critical barrier to agricultural and food security assessments and operational rapid response to humanitarian crises. These data are especially lacking in regions dominated by smallholder agriculture that are most vulnerable to food insecurity (e.g., sub-saharan Africa).

Machine learning techniques have commonly been used to identify cropland in EO data. The majority of studies use tree-based classifiers (primarily random forests or decision trees) or neural network/deep learning methods (primarily recurrent or convolutional neural networks). However, prior methods have typically been applied to small and/or spatially homogeneous areas [14, 16, 27], require large training datasets (hundreds of thousands of labels) [20, 22, 23, 32], or use EO inputs with insufficient spatial resolution for reliably detecting smallholder farms which often have field sizes <1 ha [4, 17, 24, 30, 31, 33]. In many regions, particularly developing countries and regions dominated by smallholder farming, the lack of readily accessible ground truth data is a major barrier for training and validating machine learning techniques for cropland classification [5].

In this paper, we demonstrate a novel method for rapidly generating cropland maps for areas with little to no available labeled data over a large heterogeneous area. This is achieved by leveraging global and local datasets. The global dataset contains tens of thousands of crowdsourced labels from diverse geographies [26] to learn general features for identifying cropland globally. These global examples supplement a much smaller dataset of curated,

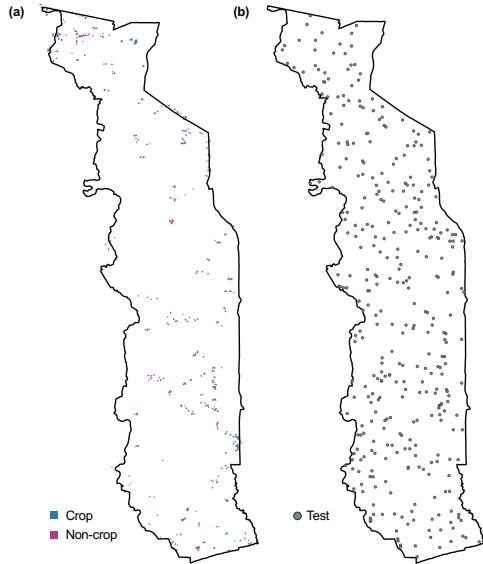


Figure 1: Distribution of labeled examples in training (a) and test (b) datasets.

local examples (manually labeled using photointerpretation of high-resolution satellite imagery) to learn features specific to cropland in the target mapping region. Combined, these datasets provide sufficient examples to train a Long Short Term Memory network (LSTM) [11] to predict the presence of cropland in each pixel given a monthly time-series of Sentinel-2 multispectral observations.

The COVID-19 global pandemic has severely impacted the global food system, and the threat of food shortages and rising food prices has required countries at risk of food insecurity to rapidly design and enact aid programs in response. In the West African country of Togo, the government announced a program to boost national food production in response to the COVID-19 crisis by distributing aid to farmers, but found that high-resolution spatial information about the distribution of farms across Togo (which primarily consist of smallholder farms under 1 ha) was a critical piece of missing information for designing this program. Using our method, we were able to generate an accurate high-resolution (10 m) map of cropland in Togo for 2019 in under 10 days to respond to the Togolese government's immediate need for information about the distribution of farms across the country. We showed that our model outperforms existing publicly-available cropland maps for Togo [4, 33] while also being higher resolution and more recent. We have delivered this product to the Togolese government where it is being used to help inform decision-making in response to the COVID-19 pandemic, thus demonstrating a successful transition of machine learning research to operational rapid response for a real humanitarian crisis. We have made our map as well as all data and code used in this study publicly available to facilitate future research and operational systems in data-sparse regions.

2 METHODS

2.1 Labeled Data

2.1.1 Crowdsourced Labels. We leveraged a large (35,866 labels) crowd-sourced dataset with labels of “crop” or “non-crop” from diverse, globally-distributed locations [26] to help the model learn features that are useful for detecting a wide range of cropland globally. Details about sample selection, labeling procedures, and quality assessment/control measures can be found in Bayas et al. [1]. Each location was labelled by multiple labelers; to turn these labels into binary labels, we took the mean of the labels for each location, with a threshold of 0.5 to define a point as “crop” or “non-crop.”

2.1.2 Active Labeling. To generate training labels, experts in photointerpretation of agricultural land cover (authors Becker-Reshef, Barker, Hosseini, Kerner, Munshell, and Paliyam) drew polygons over pixels with labels of “crop” or “non-crop.” We used a combination of basemaps for photointerpretation: a SkySat 72cm composite of January-March 2019 [29], PlanetScope composites of April-July and July-October 2019 [29], and Google Earth Pro basemaps (comprised primarily of high-resolution Maxar images). We used QGIS and Google Earth Pro to draw polygons over basemaps. The locations of labels were chosen to be distributed across different agroecological zones and non-crop land cover types observed across Togo. The central pixel within each polygon was used for training. We obtained 394 crop and 194 non-crop labels for the initial training dataset to supplement the 43 examples within Togo in the See [26] dataset (631 examples total). After training the model with the hand-labeled and See [26] dataset, we collected additional labeled examples (294 crop and 394 non-crop) by analyzing the predicted map and identifying areas of model confusion (e.g., transient vegetation on shorelines that the model might confuse with crops). This allowed us to maximize the utility of each training example, as it focused on places where the model was weakest rather than adding redundancy to the dataset. This resulted in a final training dataset of 1,319 hand-labeled examples, visualized spatially in Figure 1a.

2.2 EO Data

We used Copernicus Sentinel-2 surface reflectance (Level 2A) observations that corresponded to label locations as input to the model. We exported $160 \text{ m} \times 160 \text{ m}$ patches for each labeled pixel using Google Earth Engine (GEE) [9], then extracted the closest pixel within the patch to the label location. To construct a cloud-free timeseries representation of each pixel, we used the algorithm from Schmitt et al. [25], which finds the least cloudy pixel within a defined time period. We used 12 30-day time periods to construct a monthly timeseries spanning 360 days for each labeled pixel. We used all spectral bands except for B1 (coastal aerosol) and B10 (cirrus SWIR), and added the normalized difference vegetation index (NDVI), computed as $\text{NDVI} = \frac{B08 - B04}{B08 + B04}$. All bands (which range from 10 to 60m resolution) were upsampled to 10 m during GEE export. We used observations acquired between March of Year N and March of Year N+1 where N is the year the labels were created for. Thus for See [26], we used observations acquired March 2017–March 2018 and for our hand-labeled dataset we used observations acquired March 2019–March 2020.

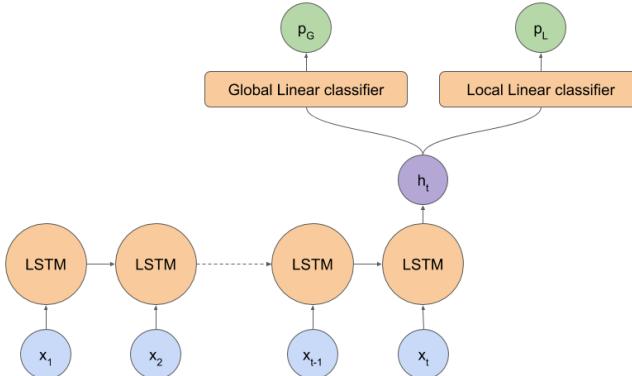


Figure 2: A multi-headed LSTM based model for pixel-wise crop identification. The model has two classification layers: a global linear classifier to classify instances outside Togo, and a local linear classifier to classify instances within Togo.

2.3 Model

We trained a one layer multi-headed LSTM model (Figure 2) to predict whether a pixel contained cropland (1) or not (0). The input was expressed as a timeseries $X = \{x_1, x_2, \dots, x_{12}\}$, where each timestep $x_i \in \mathbb{R}^{1 \times 14}$ consisted of the least-cloudy composite over a 30-day period of 11 optical bands (all bands except B01 and B10) plus NDVI. The LSTM hidden layer had 64 units. A classifier and a sigmoid activation were used on the final hidden output to map the hidden layer output (logits) to a value that can be interpreted as the posterior probability of crop presence in the pixel.

We performed a grid search that optimized AUC ROC on the validation set to choose values for two hyperparameters: number of linear layers in the classifier (1 or 2) and whether or not a dropout of 0.2 should be applied between each timestep in the LSTM.

To allow the model to focus on examples within Togo while also learning from the global See [26] dataset, we employed a multi-headed LSTM model, where one classifier was used for all examples within Togo (local instances) and another was used for all other examples (global instances). All instances were shuffled during training: a batch could therefore contain a combination of global and local data instances. To combine the results from both classifiers, we used the following loss function to train the model:

$$\mathcal{L} = \frac{W}{\alpha} \mathcal{L}_{\text{global}} + \mathcal{L}_{\text{local}} \quad (1)$$

Where $\mathcal{L}_{\text{global}}$ and $\mathcal{L}_{\text{local}}$ are the binary cross-entropy losses for the global and local data instances respectively, and α is a weighting parameter. W is a batch-specific value, used to weight the ratio of global and local instances in a batch:

$$W = \frac{\text{Number of global instances in batch}}{\text{Number of local instances in batch}} \quad (2)$$

In our experiments, we used $\alpha = 10$. We used the Adam optimizer with default parameters to tune weights during training [12]. To determine when to stop training, we used early stopping with a patience of 10 using a validation set consisting of 20% of the combined See [26] and hand-labeled Togo datasets. We implemented the model and experiments using PyTorch [18] in python. All code

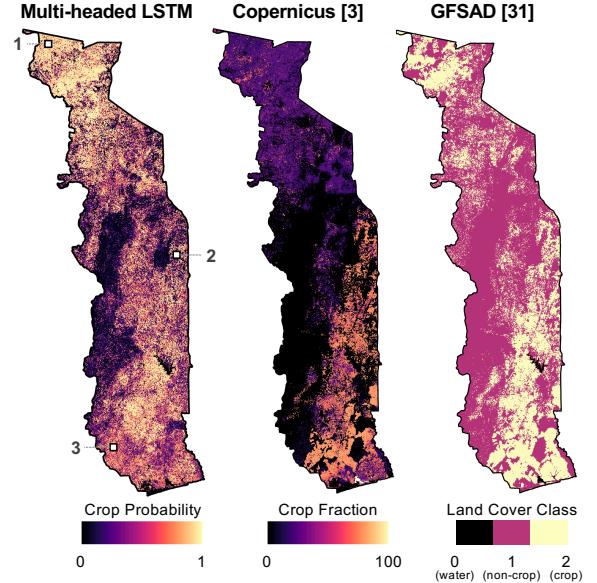


Figure 3: Cropland maps of Togo generated using our method, Buchhorn et al. [4], and Xiong et al. [33]. White boxes indicate locations of insets shown in Figure 4.

is publicly available at <https://github.com/nasaharvest/togo-crop-mask>.

3 RESULTS

We selected points for a test dataset by randomly sampling 350 points within the Togo country boundaries, constrained by a minimum distance of 50 m between samples (Figure 1b). Four experts (authors Barker, Kerner, Munshell, and Nakalembe) labeled each example. Labels were determined by majority vote in order to mitigate label noise in the absence of groundtruth validation, and ties were discarded. The consensus test dataset contained 106 crop and 200 non-crop examples (306 total test examples).

We compared two experimental setups to evaluate the contribution of the See [26] dataset:

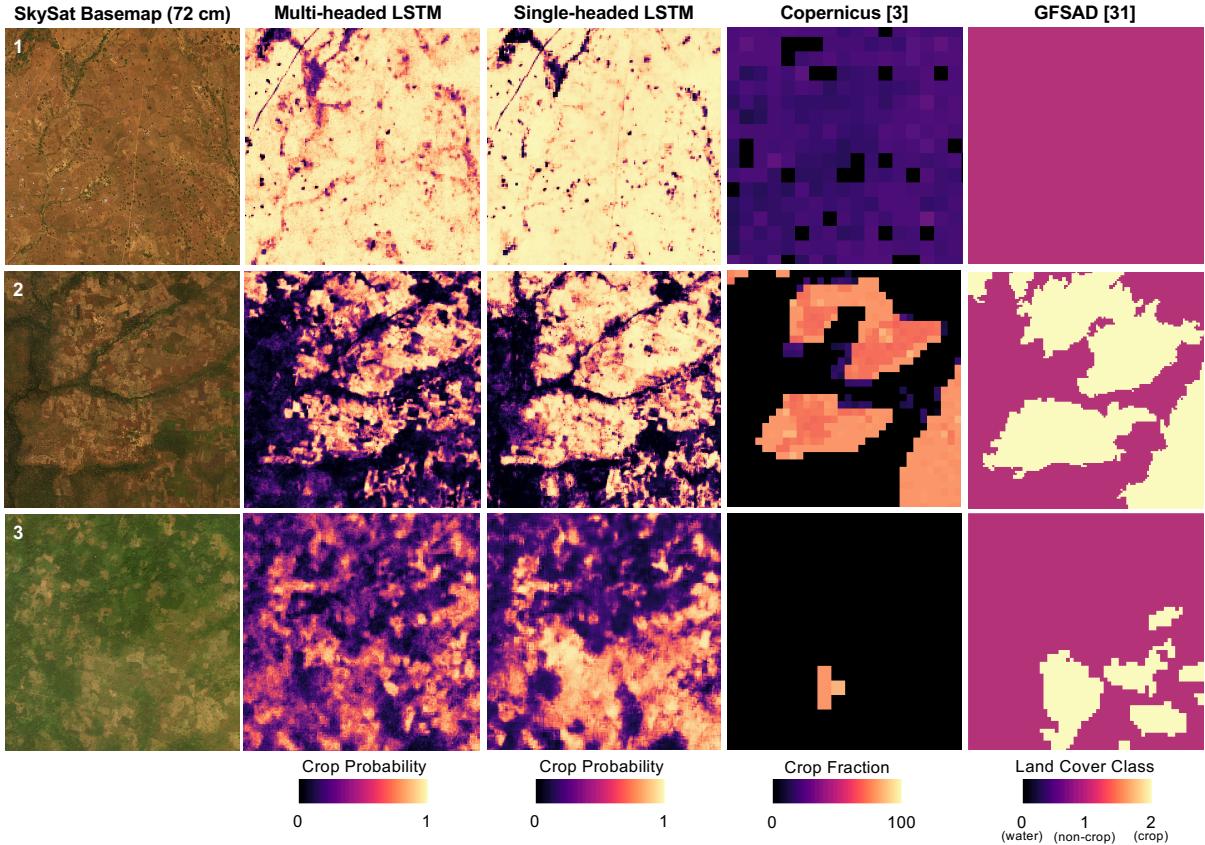
- A single-headed LSTM trained on only the hand-labeled Togo data
- A multi-headed LSTM trained on both the hand-labeled Togo and See [26] data

In both cases, we used a model with hyperparameters determined by the grid search described in 2.3. For the single headed model, this consisted of a single classification layer in the classifier *without* dropout in the LSTM. For the multi-headed model, this consisted of two classification layers in the classifier *with* dropout in the LSTM.

We compared our results to two publicly-available cropland maps of Togo. One map is Global Food Security-support Analysis Data (GFSAD) cropland extent map of Africa [33]. This map has 30 m resolution and was produced using two pixel-based supervised classifiers, a Random Forest and a Support Vector Machine (SVM), and one object-based classifier, a Recursive Hierarchical Image Segmentation classifier. The GFSAD map does not provide cropland

Table 1: Results for each method evaluated on the test set. The best results are in bold.

Description	Accuracy	AUC	Precision	Recall	F1
Buchhorn et al. [4] (Copernicus 100m)	0.67	0.74	0.54	0.24	0.33
Single-headed LSTM (hand-labeled only)	0.71	0.90	0.59	0.93	0.72
Xiong et al. [33] (GFSAD 30m)	0.74	-	0.62	0.62	0.62
Support Vector Machine (hand-labeled only)	0.79	0.88	0.65	0.84	0.73
Random Forest (hand-labeled only)	0.81	0.90	0.66	0.88	0.76
Multi-headed LSTM (hand-labeled + See [26])	0.83	0.91	0.81	0.68	0.74

**Figure 4: Qualitative comparison of cropland maps produced using our multi-headed and single-headed LSTM methods, Buchhorn et al. [4], and Xiong et al. [33] for three example regions, with SkySat 72 cm basemap (April-March composite) for reference. Locations of each inset indicated by white boxes in Figure 3.**

probabilities, but a discrete class assigned to each pixel of 0 (water), 1 (non-cropland), or 2 (cropland). The most recent GFSAD product is from 2015. Prior to our map, the GFSAD map is to our knowledge the highest resolution cropland map publicly available for Togo. Figure 3 (right) shows the GFSAD land cover map for Togo. The second map we compared our results to is the Copernicus Land Cover cropland map [4]. The Copernicus map specifies the fraction of each pixel covered by crops (using values 0-100) at 100 m resolution and was produced using a combination of random forests and expert rules with satellite data inputs from the PROBA-V sensor. The most recent Copernicus map is for 2018. To enable quantitative

comparison with our results, we treated crop fraction as analogous to crop probability and computed all performance metrics using our test dataset (Table 1); however, we emphasize that this may not be how the map was intended to be used and qualitative comparisons may be more faithful to the intended use of the Buchhorn et al. [4] map. Figure 3 (center) shows the Copernicus crop fraction map for Togo. A direct comparison between our method and Xiong et al. [33] and Buchhorn et al. [4] would require running all three methods on the same data (i.e., 10m Sentinel-2 observations paired with our labeled examples). Since neither the code for these methods nor sufficient information for reproducing them faithfully are available,

we additionally performed two baseline comparisons to Random Forest and SVM classifiers which are the basis for the Xiong et al. [33] and Buchhorn et al. [4] methods. For the Random Forest, we used 100 estimators with no maximum depth (nodes were expanded until all leaves were pure, or until the leaves contained fewer than 2 samples). For the SVM, we used a radial basis function kernel. We implemented both methods using Scikit-learn in python [19].

To assess performance, we computed the area under the receiver operating characteristic (ROC) curve (AUC score), accuracy (0.5 threshold on posterior probabilities), precision (user's accuracy), recall (producer's accuracy), and F1-score (harmonic mean between precision and recall). We did not compute AUC for GFSAD, which only provides the discrete predicted class. The results are reported in Table 1. Figure 3 (left) shows our predicted cropland map.

To qualitatively compare our method compared with Xiong et al. [33] and Buchhorn et al. [4], we show each map compared to the SkySat 72cm basemap (true color) for several example locations in Figure 4. Each region in Figure 4 contains primarily cropland, though the level of vegetation in non-crop areas varies in each image since Togo's climate ranges from dry savanna in the north to tropical in the south. While all maps detect most fields in the second row, the Xiong et al. [33] and Buchhorn et al. [4] maps failed to detect most of the fields in the first and third rows.

Our resulting map of cropland probabilities at 10 m resolution as well as all hand-labeled training and testing labels are publicly available at: <https://doi.org/10.5281/zenodo.3836629>.

4 DISCUSSION

The multi-headed model trained on the combined hand-labeled and See [26] dataset had the highest score for most performance metrics. While AUC score is a measure of model performance across a range of possible thresholds on posterior probability, the decision boundary when using the sigmoid function as the output activation layer is 0.5. While in practice a different threshold could be chosen, it is not clear how to construct a large and diverse enough validation set that could be used to reliably select an alternate threshold over very large, heterogeneous regions (e.g., the country of Togo). Thus, the improved accuracy and precision of the multi-headed model (using a threshold of 0.5) - in addition to its improved AUC score - made it preferable. Additionally, since the multi-headed model included training data from prior years (2017) and global instances of cropland [26], we expect that this model would generalize better to potential domain shift in future years than the single-headed model trained with only one year of local data. Figure 4 shows that the multi-headed LSTM and single-headed LSTM have similar patterns of detection, but overall the multi-headed LSTM has less extreme probabilities (values closer to 0.5 than 0 or 1). This suggests that the global examples used to train the multi-headed LSTM act like a regularizer for the classifier by moderating the confidence of predictions and preventing overfitting.

We found that the Buchhorn et al. [4] map had the lowest performance for the test dataset, with an F1 score significantly lower than other methods, though we note that these results could underestimate the performance because the map reports the fraction of each pixel covered by crops, not probability, in each pixel. This lower performance may, in part, be due to the coarse spatial resolution

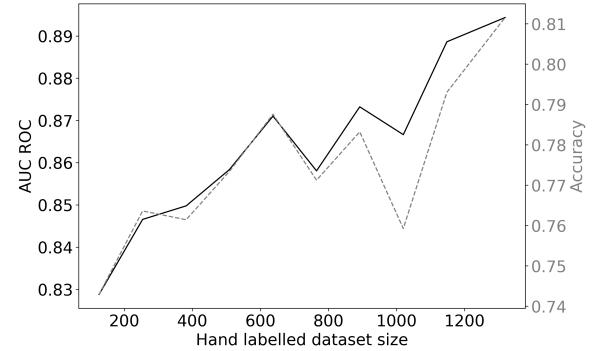


Figure 5: A plot of AUC ROC and accuracy scores of the multi-headed model as the size of the hand-labeled dataset size increases. Each score is calculated from a mean of 3 runs with different random seeds.

of EO data used to produce the map and the growth in cropland area that has likely occurred since the map was produced in 2015. Similarly, the lower performance of the Xiong et al. [33] could also, in part, be due to lower spatial resolution of input data (30 m) and changes in cropland area that might have occurred between 2018 (when the most recent Xiong et al. [33] map was produced) and 2019.

4.1 Sensitivity to number of local examples

To assess the sensitivity of the model to the frequency and number of local (Togo) examples in the training dataset and provide a guide for future labeling campaigns, we measured the performance of the multi-headed model on the test set as a function of the number of hand-labeled examples used in training (Figure 5). Although the performance of the model increased as the size of the hand-labeled dataset increased, the model performs well even with very little labeled data. This suggests this technique may be applied to other data-sparse regions with a comparable or fewer number of labels as we created for this study.

4.2 Consensus in photointerpreted labels

As described in Section 3, we created a test dataset by randomly sampling 350 points within Togo and having four experts label each point as "crop" or "non-crop" using photointerpretation of high-resolution satellite imagery. For each point, we chose the label assigned by the majority of labelers (ties were discarded). Figure 6 shows the number of examples for which each pair of labelers chose the same label. Agreement between pairs of labelers ranged from 229-278 (65%-79%) of the total points and all labelers agreed unanimously on 181 (52%) of the points. This level of consensus was surprisingly low, and while our photointerpretation-based dataset is a high-quality surrogate for ground-truth labels, it underscores the difficulty for even experts to determine land cover types from high-resolution satellite imagery, particularly in developing countries. This should be considered in the design of future photointerpretation-based labeling efforts, especially crowdsourced efforts in which labelers may not be experts. We plan to conduct a future study to assess agreement between labels determined by

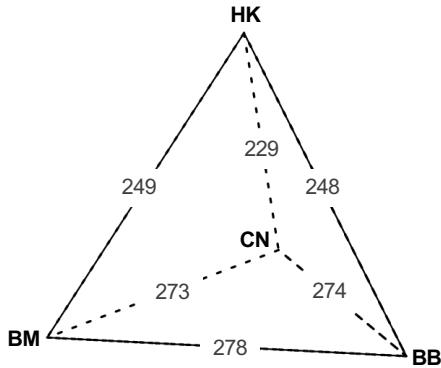


Figure 6: Tetrahedron showing number of examples (of 350 total) for which each pair of labelers (indicated by initials) agreed.

experts for the same points from photointerpretation vs. field observation to better characterize the level of accuracy to expect from photointerpreted labels for different regions and land cover types.

4.3 Lessons learned for rapid response

The urgent need for a cropland map of Togo that captured small-holder farms (usually <1 ha) required us to leverage our research methods to rapidly deliver a result for operational use in less than 10 days. During this process, we gathered several “lessons learned” that we share to help prepare other researchers to leverage their methods for rapid humanitarian response in the future, as well as encourage community practices that facilitate rapid response.

- Good software engineering practices (such as developing modular and well-documented code) is often prioritized in operational rather than research settings. However, developing research methods motivated by their use for rapid humanitarian response in the future requires integration of operational practices into research workflows. We found that developing research methods using operational software engineering practices enabled us to quickly execute experiments and apply the method to large geographic areas when the need for rapid response arose.
- Use of cloud services (in this case AWS) allowed us to easily scale our model (particularly during country-wide inference) and enable faster computation in a short period of time. Additionally, the use of EC2 instances enabled processes to be run, monitored, and modified by a distributed team rather than a single user.
- File transfer/download remains a bottleneck for rapid response workflows. Transferring data from storage (e.g., via Google Earth Engine, the Sentinel-2 AWS S3 bucket, or USGS Earth Explorer) to compute (e.g., AWS) constituted a significant part of the time to deliver our cropland map. There is a community need for better tools to enable searching, querying, and subsetting of geospatial data on cloud platforms.
- Given the complexity of deep learning architectures such as Inception-v3 [28] or ResNet [10] designed for image classification, it might seem that classification of land cover or crops

in remote sensing data—which often exhibit high intraclass variance and low inter-class variance—would also require complex architectures. However, we found that a relatively simple, shallow architecture achieved good results for this task, and has the benefit of being faster in training/inference, requiring fewer examples, and being less susceptible to overfitting than more complex architectures.

- The GFSAD [33] and Copernicus [4] cropland maps were easy to find and download and were well-documented, which enabled us to quickly benchmark our map against existing maps. To give others this same benefit with our cropland map, we have made our cropland map, code, and training labels publicly available.

5 CONCLUSION

Cropland maps are critical inputs for decision-makers to rapidly design and enact policies during humanitarian crises, e.g., to deliver aid or boost productivity for farmers, yet unavailability of high-resolution, up-to-date cropland maps in most countries globally precludes their use for rapid response. The development of cropland maps, in turn, is limited by the availability of ground data that can be used for training and validating machine learning classifiers.

We present a method for cropland classification in regions with little to no available ground data that uses a multi-headed LSTM network to learn global and local features for identifying cropland based on multispectral time-series Sentinel-2 observations. We used this method to create a cropland map of Togo at 10 m resolution in fewer than 10 days to assist the Togolese government in decision-making about aid distribution to farmers during the COVID-19 crisis. This demonstrated the successful transition of machine learning research to operational use in rapid response for a real humanitarian crisis, and we provided lessons learned from this experience to facilitate this transition for other researchers in the future. Additionally, our approach gives an example of how to leverage existing readily-available labeled datasets and reduce the number of data points to be collected in field campaigns, which can be costly as well as infeasible due to regional insecurity or travel restrictions (e.g., during the COVID-19 pandemic). In future work, we plan to investigate leveraging our multi-headed LSTM approach for crop type classification, subsetting the Geo-Wiki dataset based on geographical or agroecological zones, and including Sentinel-1 synthetic aperture radar data in the model input to mitigate the effects of clouds during the growing season. Additionally, we will apply our method for cropland mapping in other data-sparse African countries.

ACKNOWLEDGMENTS

We would like to thank Tanya Harrison, Zara Khan, and Charlie Candy from Planet, Inc., for making SkySat and PlanetScope basemaps available for this study. We are also grateful to Noel Gorelick from Google Earth Engine for helping to accelerate our Google Earth Engine processing workflow in order to meet this rapid response timeline. Finally, we would like to thank Steffen Fritz, Linda See, and the rest of the Geo-Wiki team for providing the publicly-available, global, crowdsourced cropland dataset critical to our method.

REFERENCES

- [1] Juan Carlos Laso Bayas, Myroslava Lesiv, François Waldner, Anne Schucknecht, Martina Duerauer, Linda See, Steffen Fritz, Dilek Fraisl, Inian Moorthy, Ian McCallum, et al. 2017. A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. *Scientific data* 4 (2017), 170136.
- [2] Inbal Becker-Reshef, Christina Jade Justice, Brian Barker, Michael Laurence Humber, Felix Rembold, Rogerio Bonifacio, Mario Zappacosta, Mike Budde, Tamuka Magadzire, Chris Shitote, Jonathan Pound, Alessandr Constantino, Catherine Nakalembe, Kenneth Mwangi, Shinichi Sobue, Terence Newby, Alyssa Whitcraft, Ian Jarvis, and James Verdin. 2020. Strengthening agricultural decisions in countries at risk of food insecurity: The GEOGLAM Crop Monitor for Early Warning. *Remote Sensing of Environment* 237 (2020), 111553.
- [3] Inbal Becker-Reshef, Eric Vermote, Mark Lindeman, and Christopher Justice. 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote sensing of environment* 114, 6 (2010), 1312–1323. <https://doi.org/10.1016/j.rse.2010.01.010>
- [4] Marcel Buchhorn, Bruno Smets, Luc Bertels, Myroslava Lesiv, Nandin-Erdene Tsendsazar, Martin Herold, and Steffen Fritz. 2019. *Copernicus Global Land Service: Land Cover 100m: epoch 2018: Africa demo*. <https://doi.org/10.5281/zenodo.3518087>
- [5] Sylvain Coutu, Inbal Becker-Reshef, Alyssa Whitcraft, and Chris Justice. 2020. Food security: underpin with public and private data sharing. *Nature* 578, 7796 (2020), 515. <https://doi.org/10.1038/d41586-020-00241-y>
- [6] Andrew Dorward and Ephraim Chirwa. 2011. The Malawi agricultural input subsidy programme: 2005/06 to 2008/09. *International Journal of Agricultural Sustainability* 9, 1 (2011), 232–247. <https://doi.org/10.3763/ijas.2010.0567>
- [7] Belen Franch, Eric Vermote, Inbal Becker-Reshef, Martin Claverie, Jianxi Huang, Jinshui Zhang, Chris Justice, and Jose Antonio Sobrino. 2015. Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information. *Remote Sensing of Environment* 161 (2015), 131–148. <https://doi.org/10.1016/j.rse.2015.02.014>
- [8] Belen Franch, Eric Vermote, Jean-Claude Roger, Emilie Murphy, Inbal Becker-Reshef, Chris Justice, Martin Claverie, Jyoteshwar Nagol, Ivan Csizsar, Dave Meyer, et al. 2017. A 30+ year AVHRR land surface reflectance climate data record and its application to wheat yield monitoring. *Remote Sensing* 9, 3 (2017), 296. <https://doi.org/10.3390/rs9030296>
- [9] Noel Gorelick, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202 (2017), 18 – 27. <https://doi.org/10.1016/j.rse.2017.06.031>
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- [11] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-term Memory. *Neural computation* 9 (12 1997), 1735–80. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [12] Diederik Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [13] David Lobell. 2013. The use of satellite data for crop yield gap analysis. *Field Crops Research* 143 (2013), 56–64. <https://doi.org/10.1016/j.fcr.2012.08.008>
- [14] Dinh Ho Tong Minh, Dino Ienco, Raffaele Gaetano, Nathalie Lalande, Emile Ndikumana, Faycal Osman, and Pierre Maurel. 2018. Deep recurrent neural networks for winter vegetation quality mapping via multitemporal SAR Sentinel-1. *IEEE Geoscience and Remote Sensing Letters* 15, 3 (2018), 464–468. <https://doi.org/10.1109/LGRS.2018.2794581>
- [15] Engineering National Academies of Sciences and Medicine. 2019. *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space: An Overview for Decision Makers and the Public*. The National Academies Press. <https://doi.org/10.17226/25437>
- [16] Emile Ndikumana, Nicolas Baghdadi, Dominique Courault, Laure Hossard, and Dinh Ho Tong Minh. 2018. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing* 10 (08 2018), 1217. <https://doi.org/10.3390/rs10081217>
- [17] Adam Oliphant, Prasad Thenkabail, Pardhasaradhi Teluguntla, Jun Xiong, Murali Krishna Gumma, Russell Congalton, and Kamini Yadav. 2019. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud. *International Journal of Applied Earth Observation and Geoinformation* 81 (2019), 110–124.
- [18] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Razis, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32*, 8026–8037. <http://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- [19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
- [20] Charlotte Pelletier, Silvia Valero, Jordi Inglada, Nicolas Champion, and Gérard Dedieu. 2016. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment* 187 (2016), 156–168. <https://doi.org/10.1016/j.rse.2016.10.010>
- [21] John Porter, Liyong Xie, Andrew Challinor, Kevern Cochrane, S Mark Howden, Muhammad Mohsin Iqbal, Maria Travasso, VR Barros, CB Field, DJ Dokken, et al. 2017. Food security and food production systems. (2017). <https://hdl.handle.net/10568/68162>
- [22] Marc Russwurm and Marco Körner. 2017. Temporal Vegetation Modelling Using Long Short-Term Memory Networks for Crop Identification From Medium-Resolution Multi-Spectral Satellite Images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- [23] Marc Rußwurm, Charlotte Pelletier, Maximilian Zollner, Sébastien Lefèvre, and Marco Körner. 2020. BreizhCrops: A Time Series Dataset for Crop Type Mapping. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences ISPRS* (2020) (2020).
- [24] Kaboro Samasse, Niall Hanan, Julius Anchang, and Yacouba Diallo. 2020. A High-Resolution Cropland Map for the West African Sahel Based on High-Density Training Data, Google Earth Engine, and Locally Optimized Machine Learning. *Remote Sensing* 12, 9 (2020), 1436.
- [25] Michael Schmitt, Lloyd Hughes, Chunping Qiu, and Xiao Zhu. 2019. Aggregating cloud-free Sentinel-2 images with Google Earth Engine. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences IV-2/W7* (09 2019), 145–152. <https://doi.org/10.5194/isprs-annals-IV-2-W7-145-2019>
- [26] Linda See. 2017. A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. <https://doi.org/10.1594/PANGAEA.873912>
- [27] Sergii Skakun, Christopher Justice, Natalia Kussul, Andrii Shelestov, and Mykola Lavreniuk. 2019. Satellite data reveal cropland losses in South-Eastern Ukraine under military conflict. *Frontiers in Earth Science* 7 (2019), 305.
- [28] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2818–2826.
- [29] Planet Team. 2018-. Planet Application Program Interface: In Space for Life on Earth. <https://api.planet.com>
- [30] Pardhasaradhi Teluguntla, Prasad Thenkabail, Adam Oliphant, Jun Xiong, Murali Krishna Gumma, Russell Congalton, Kamini Yadav, and Alfredo Huete. 2018. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing* 144 (2018), 325–340.
- [31] François Waldner, Matthew C Hansen, Peter V Potapov, Fabian Löw, Terence Newby, Stefanus Ferreira, and Pierre Defourny. 2017. National-scale cropland mapping based on spectral-temporal features and outdated land cover information. *PLoS one* 12, 8 (2017).
- [32] Sherrie Wang, William Chen, Sang Michael Xie, George Azzari, and David Lobell. 2020. Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery. *Remote Sensing* 12, 2 (2020), 207.
- [33] Jun Xiong, Prasad Thenkabail, J Tilton, Murali Krishna Gumma, Pardhasaradhi Teluguntla, Russell Congalton, Kamini Yadav, J Dungan, Adam Oliphant, J Poehnelt, C Smith, and R Massey. 2017. *NASA Making Earth System Data Records for Use in Research Environments (MEaSUREs) Global Food Security-support Analysis Data (GFSAD) Cropland Extent 2015 Africa 30 m V001*. [https://doi.org/10.5067/MEAASUREs/GFSAD/GFSAD30AFCE.001](https://doi.org/10.5067/MEASUREs/GFSAD/GFSAD30AFCE.001)