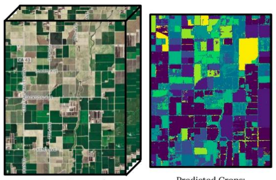
Crop Type Mapping with Multi-Temporal and Multi-Spatial Satellite Imagery

Rose Rustowicz, Robin Cheong, and Lijing Wang

Overview

We combine remote sensing imagery with deep learning algorithms to distinguish and map crop type from space.

Visualization of the objective: Use temporal satellite imagery to map crop type from space.



Input Temporal Image Stack

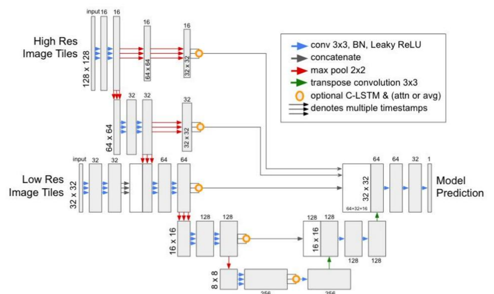
Predicted Crops: color corresponds to different crop types

Motivation: Over 800 million people in the world are undernourished and 80% of consumed food in the developing world comes from smallholder farms [1]. Having a better understanding of smallholder farms via crop type maps can provide unprecedented insight into food systems and food security. We tie this motivation into the UN's 2030 Sustainable Development Goals to contribute to goal three: zero hunger.

We focus in South Sudan and Ghana, Africa, where food security is of particular importance. We encounter challenges with sparse data labels, class imbalance, and high cloud cover

[1] https://www.un.org/sustainabledevelopment/hunger/

Model Architecture



Data

Satellite Sources

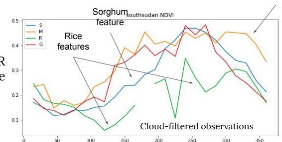
Sentinel-1 (S1) and Sentinel-2 (S2)

- Spatial resolution: 10 m
- Revisit Rate: 6-12 days
- S-1 Spectral Bands:
  - VV, VH Polarization
- S-2 Spectral Bands:
  - BGR, "Red Edge," NIR, SWIR
- Preprocessing: TOA reflectance

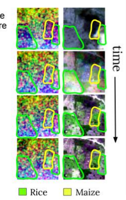
Planet

- Spatial resolution: 3m
- Revisit Rate: 1-2 days
- Spectral Bands: BGR, NIR
- Preprocessing: filtered out images with >10% cloud cover, TOA

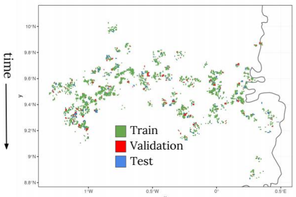
Filtered Mean Temporal Features for South Sudan Crops



Time Series Subset for S1 and S2



Ghana Dataset Splits



Ground Truth

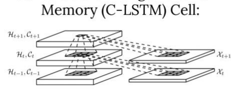
- Georeferenced polygons with crop type label

Study Regions

- South Sudan, 2017
  - 4 classes
  - ~65k pixels
- Ghana, 2017
  - 4 classes
  - ~575k pixels
- Germany, 2016
  - 17 classes
  - ~1330k pixels

Other Concepts

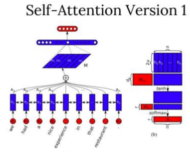
Convolutional Long Short Term Memory (C-LSTM) Cell:



$$i_t = \sigma(W_{xi} * \mathcal{X}_t + W_{hi} * \mathcal{H}_{t-1} + W_{ci} \circ \mathcal{C}_{t-1} + b_i)$$
$$f_t = \sigma(W_{xf} * \mathcal{X}_t + W_{hf} * \mathcal{H}_{t-1} + W_{cf} \circ \mathcal{C}_{t-1} + b_f)$$
$$\mathcal{C}_t = f_t \circ \mathcal{C}_{t-1} + i_t \circ \tanh(W_{xc} * \mathcal{X}_t + W_{hc} * \mathcal{H}_{t-1} + b_c)$$
$$o_t = \sigma(W_{xo} * \mathcal{X}_t + W_{ho} * \mathcal{H}_{t-1} + W_{co} \circ \mathcal{C}_t + b_o)$$
$$\mathcal{H}_t = o_t \circ \tanh(\mathcal{C}_t)$$

[2] Shi et. al., "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting", 2015. url=https://arxiv.org/pdf/1506.04214.pdf

Self-Attention Version 1



[3] Lin et. al., "A Structured Self-Attentive Sentence Embedding", *International Conference on Learning Representations (ICLR)*, 2017.

Self-Attention Version 2

Inputs: t0 → [1 x encoded features] t1 → [1 x encoded features] ... tN →
Queries: q0 → [1 x dk features] q1 → [1 x dk features] ... qN →
Keys: k0 → [1 x dk features] k1 → [1 x dk features] ... kN →
Values: v0 → [1 x dv features] v1 → [1 x dv features] ... vN →

Scores t0: s0 → <q0, k0> s1 → <q0, k1>
Scores t1: s0 → <q1, k0> s1 → <q1, k1>
....
Scores tN

Wq = [enc feats x dk]
Wk = [enc feats x dk]
Wv = [enc feats x dv]

Final values: fv0 → v0 * max score t0 fv1 → v1 * max score t1

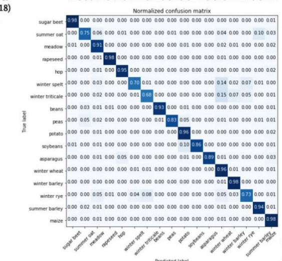
[4] Vaswani et. al., "Attention is All you Need," 31st Conference on Neural Information Processing Systems (NIPS 2017).

Experiments

Method vs. State of the Art on Large Germany Data

Crop type	Our Model TestF1	Rußwurm & Körner (2018) Test F1
Sugar Beet	0.951	0.853
Summer Oat	**0.806**	**0.758**
Meadow	0.908	0.882
Rapeseed	**0.977**	**0.926**
Hop	0.942	0.917
Winter Spelt	**0.719**	**0.656**
Winter Triticale	0.714	0.618
Beans	**0.950**	**0.896**
Peas	0.885	0.804
Potato	**0.944**	**0.895**
Soybeans	0.895	0.877
Asparagus	**0.897**	**0.837**
Winter Wheat	0.962	0.903
Winter Barley	**0.961**	**0.910**
Winter Rye	0.732	0.607
Summer Barley	**0.934**	**0.854**
Maize	0.979	0.939
Macro Avg. f1-score	**0.892**	**0.831**
Overall Accuracy	**0.949**	**0.897**
Kappa Coefficient	**0.937**	**0.870**

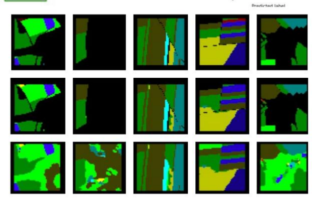
Test Performance Confusion Matrix



Ground Truth

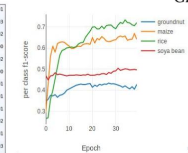
Masked Model Preds

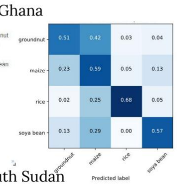
Model Preds



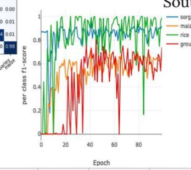
Results in Smallholder Farms in Africa

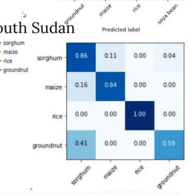
Ghana





South Sudan





	Num Pixels	Metrics	Random Forest	Our Model	Rußwurm & Körner (2018)
Ghana	~65K	accuracy macro f1	0.590 0.408	0.589 **0.571**	--
South Sudan	~575K	accuracy macro f1	0.923 **0.837**	0.842 0.775	--
Germany	~1330K	accuracy macro f1	0.677 0.251	0.949 **0.892**	0.897 0.831