Final Project Progress Report

Data-Efficient Cropland Detection from Satellite Imagery

Class: Computer Vision CSC 752 -UT1

Department of Computer Science

University of South Dakota

Group members:

Amit Kumar Patel

Jayakumar Pujar

Robin Narsingh Ranabhat

Supervisor: Dr. Lina Chato

Fall 2024

Abstract

This project explores data-efficient techniques for detecting cropland areas using satellite images. By leveraging Sentinel-2 imagery and calculating NDVI (Normalized Difference Vegetation Index) values, the project aims to identify agricultural regions. Initial models employ a simple encoder-decoder architecture, optimized for efficient pixel-level classification. Challenges include a lack of domain expertise in geospatial data and the absence of explicit training labels. This report presents early findings, methods, and proposed improvements to enhance the model's accuracy in cropland detection.

Introduction and Literature Review

Understanding and mapping cropland areas is essential for sustainable agricultural practices and resource management. Recent studies highlight the potential of machine learning and remote sensing in extracting valuable insights from geospatial data. Sentinel-2 satellite images have proven effective in various environmental monitoring applications, including vegetation and land use classification (Gorelick et al., 2017).

A prevalent approach involves calculating NDVI, which assesses vegetation health and distribution, making it a suitable metric for cropland detection. However, the literature also emphasizes the importance of annotated training data and domain-specific knowledge for accurate model performance (Vali et al., 2020). This project builds on these insights by proposing an initial encoder-decoder architecture designed to function effectively with minimal labeled data, making it highly suitable for agricultural applications in regions with limited resources.

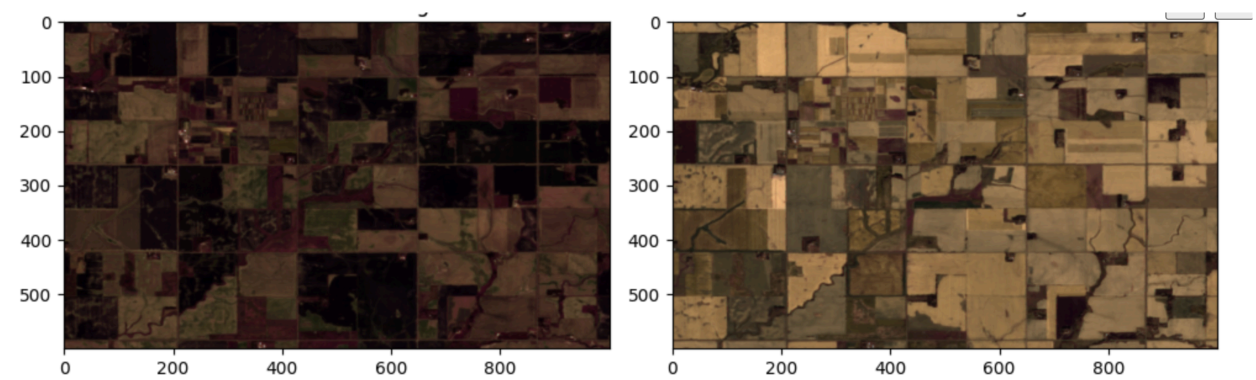
Data

The data comprises Sentinel-2 satellite imagery acquired from Google Earth Engine, focused on an agricultural region near Vermillion.

We use two primary datasets:

- **Cultivation Period:** Images taken from May to August 2023.
- **Post-Harvest Period:** Images captured from January to February 2024.

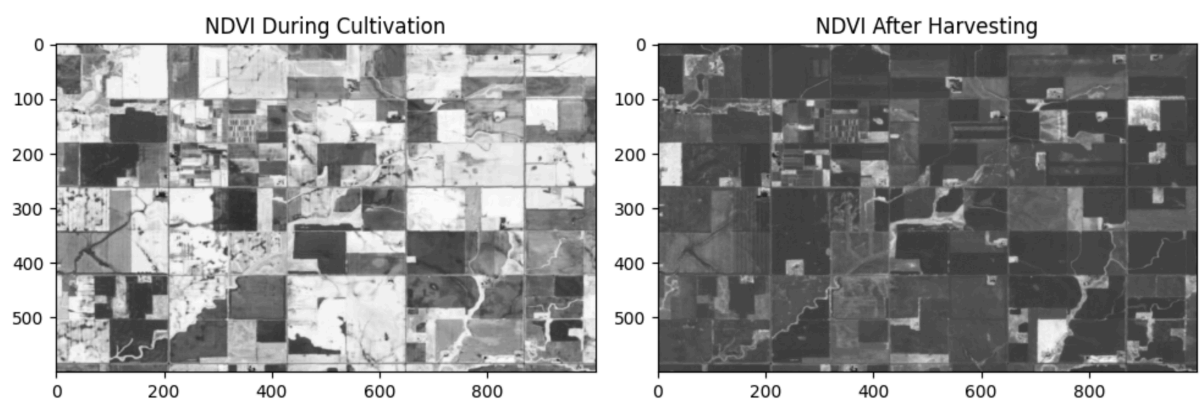
The raw data consists of two large image tif files of shape 250,000 x 250,000 x 27. 27 is the number of spectral bands captured by sentinel-2. Each Pixel represents a region of 10*10 meters square on the surface. To manage this large data, we preprocess it into 256 x 256 image chunks.



Method(s)

1. **Data Preprocessing**
   - The final tif files were generated by taking the “band-level” median of all images captured between the stated time period.
   - NDVI Calculation: We calculate NDVI for both the cultivation and post-harvest periods using the information in NIR band and Red band as follows :

     NDVI = (NIR−RED)/(NIR+RED)
     This transformation emphasizes vegetated areas.



   - Pseudo-Labeling: We create pseudo-labels by subtracting pre-harvest NDVI values from post-harvest NDVI values. The difference represents the region of cropland. This difference is mapped the output to a 0-1 scale to represent the probability of cropland presence.

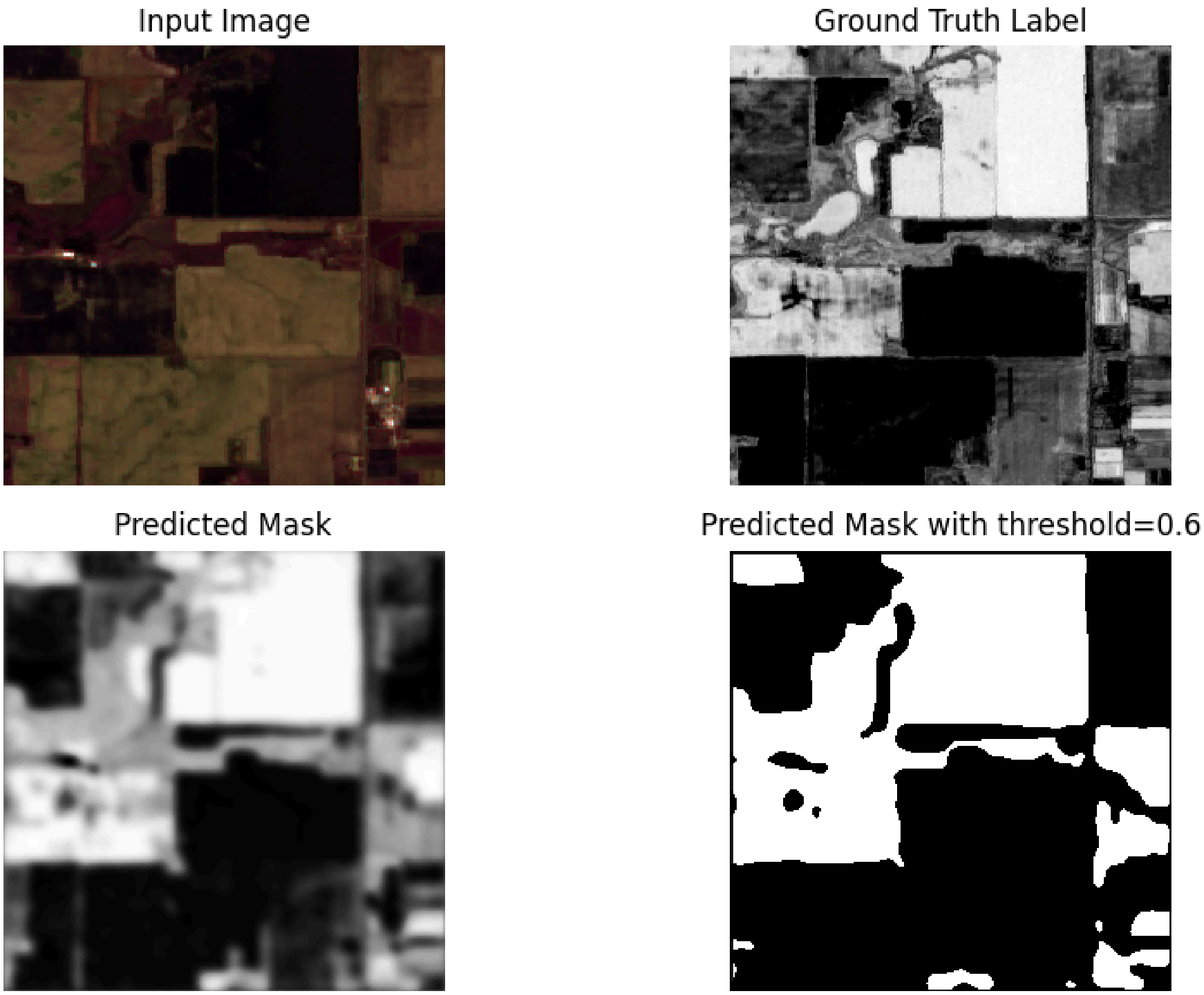
2. **Model Design**
   - Architecture: We utilize an Unet architecture to generate pixel-wise cropland probability masks. A sigmoid activation in the final layer ensures the output values fall within a 0-1 range. Below is the model architecture summary :

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1         [-1, 64, 256, 256]           1,792
              ReLU-2         [-1, 64, 256, 256]               0
            Conv2d-3         [-1, 64, 256, 256]          36,928
              ReLU-4         [-1, 64, 256, 256]               0
         MaxPool2d-5         [-1, 64, 128, 128]               0
            Conv2d-6        [-1, 128, 128, 128]          73,856
              ReLU-7        [-1, 128, 128, 128]               0
            Conv2d-8        [-1, 128, 128, 128]         147,584
              ReLU-9        [-1, 128, 128, 128]               0
         Upsample-10        [-1, 128, 256, 256]               0
           Conv2d-11         [-1, 64, 256, 256]          73,792
             ReLU-12         [-1, 64, 256, 256]               0
           Conv2d-13          [-1, 1, 256, 256]              65
          Sigmoid-14          [-1, 1, 256, 256]               0
================================================================
Total params: 334,017
Trainable params: 334,017
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.75
Forward/backward pass size (MB): 329.00
Params size (MB): 1.27
Estimated Total Size (MB): 331.02
----------------------------------------------------------------
```

Experiments and Results/Expectations

Initial Experiments:

Our preliminary model was tested on training data to assess its learning capabilities. It shows early signs of pattern recognition for cropland areas, as shown in the figure below. However, without expert-verified labels, model validation remains limited.



.

Expected Improvements:
By incorporating human-annotated labels and contrastive learning to pre-train the encoder, we expect the model's segmentation accuracy to improve between cropland and other vegetation types.

Conclusion and Discussion

Our initial model design shows potential in identifying cropland areas using limited data and without explicitly labeled training data. The lack of domain expertise and expert annotations are key challenges; however, the proposed approach of pseudo-labels is a feasible foundation for cropland detection in a data-efficient way. Further, introducing contrastive learning and pre-trained encoders should help address overfitting and enhance model generalization.

Future Work (to be completed by Week 15)

- Work on GAN models to generate post-harvest images from pre-harvest images
- Develop a small expert-annotated validation set to evaluate model accuracy.
- Fine-tune the model using contrastive learning to improve segmentation quality.
- Explore multi-task learning by predicting both cropland and general vegetation presence.

References

- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27.
- Vali, A.; Comai, S.; Matteucci, M. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. Remote Sens. 2020, 12, 2495. https://doi.org/10.3390/rs12152495
- https://code.earthengine.google.com/

Team members Contributions

All the team members were involved in data collection, pre-processing, model development, and finetuning the model. However, the main focus of the team members may be divided as follows:

Amit Kumar Patel : Data collection, domain exploration
Jayakumar Pujar : Research on Data pre-processing techniques and pseudo-labelling
Robin Narsingh Ranabhat : Model training and validation