Hansae Kim, Jinha Jung

Lyles School of Civil Engineering, kim4012@purdue.edu, Jinha@purdue.edu

Learning agricultural parcel from pre-existing cadastral parcel using unet-based deep neural network

PURDUE UNIVERSITY

ABSTRACT

Agricultural parcel plays a significant role in the agriculture industry and cropland management as it can be the foundation of crop yield estimation and crop type classification. We propose the methodology to delineate the agricultural parcel boundary using remote sensing imagery and pre-built cadastral parcel geographical information system (GIS) data. Experimental results showed that our methodology could be used to extract the agricultural parcel boundaries for a large area, which might reduce manual digitization significantly. This study illustrates the potential to generate agricultural parcel layers from pre-existing cadastral parcel layers. In addition, the proposed approach can be applied to any deep learning models using remote sensing imagery and simple GIS data with a fully automated process and can speed up the process of producing the agricultural parcel map.

INTRODUCTION

Delineating agricultural land is a cornerstone of various agriculture applications as it provides key information for crop yield estimation, crop type classification, land use management, and other practical applications.

WORKFLOW

We propose the methodology to delineate the agricultural parcel, which incorporates pre-existing cadastral parcel data, ancillary geographical information system (GIS) data, and remote sensing imagery. To this end, we first introduce the training datasets generation method. We created the label data using the cadastral parcel, road centerlines, and the normalized difference height model (NDHM) over a large area with time-series remote sensing imagery. Multitemporal training datasets are trained in the unet-based deep neural network and the trained model was then used to perform semantic segmentation of agricultural parcels. Figure 1 illustrates our general workflow.

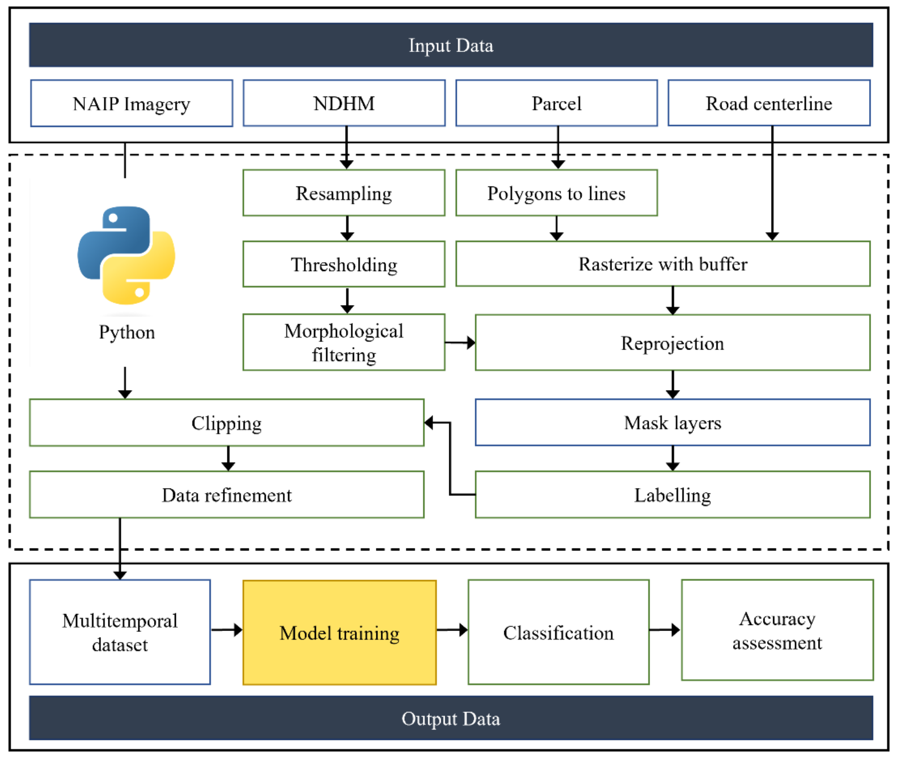

Figure 1. General workflow

TRAINING DATA GENERATION

First, the non-agricultural area was labeled as a background class using NDHM. Second, cadastral parcels and road centerlines were used to label the parcel boundary and the road classes. After creating the mask layers, label data was automatically generated. Regions of intersection between the background and the road are labeled as road, whereas regions of intersection between the background and the parcel are labeled as background. Conversely, overlapping areas between the parcel and the road are labeled as roads. Figure 2 shows examples of materials used in the study alongside generated labeled data.

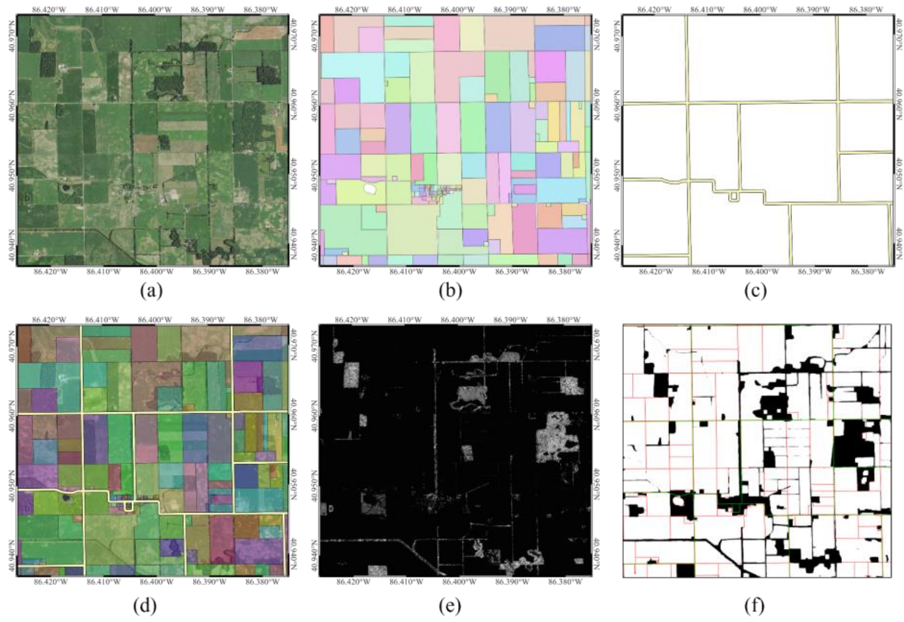

Figure 2. Training data generation

EXPERIMENTAL RESULTS

The model was trained using the generated dataset and evaluated. To mitigate the impact of inadequate representation of the dataset and to report a more robust estimate of the model's performance, the experiments were conducted in ten trials for each dataset by randomly sampling the training and validation data. Manually hand-digitized ground truth was used for a fair evaluation, assuming that the given labeled data can be noisy labeled. Figure 3 and 4 are the results using noisy dataset and clean dataset. White, black, red, and blue indicate TP, TN, FP, and FN, respectively.

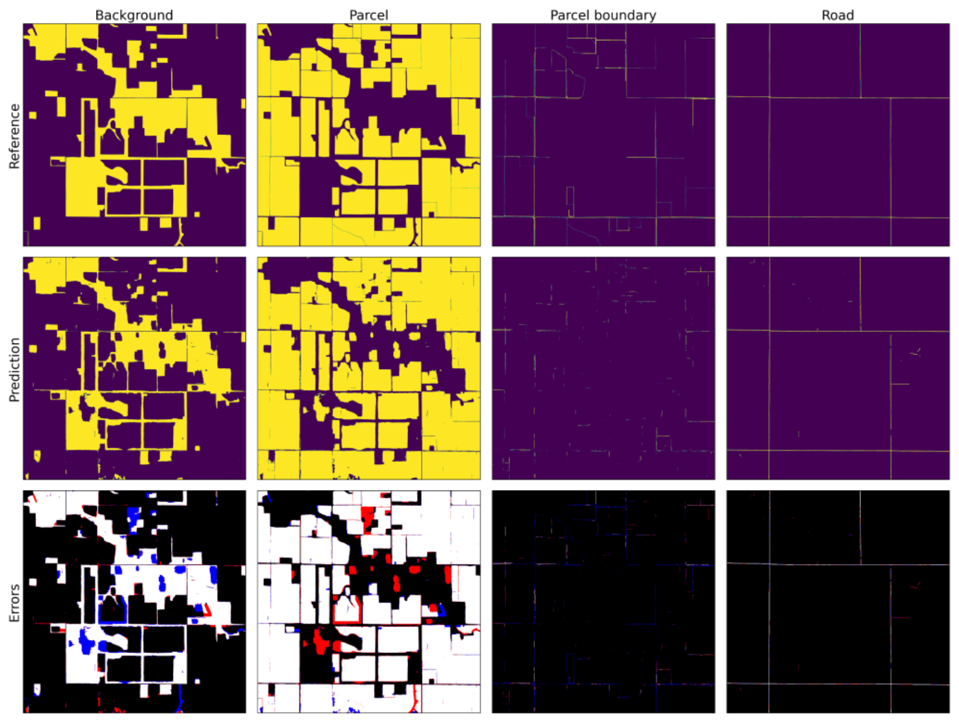

Figure 3. Results using Indiana(noisy) dataset

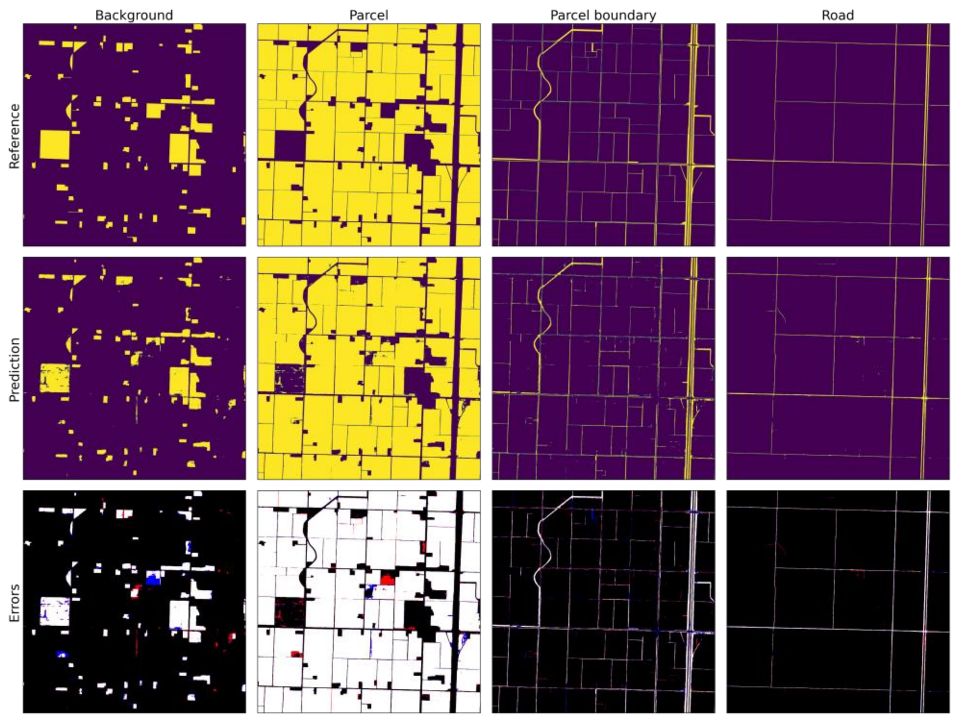

Figure 4. Results using California(clean) dataset

About our group

GDSL @ Purdue

GDSL @ Purdue University is well known for Remote Sensing, Geographic Information System (GIS), Unmanned Aircraft System (UAS), High Performance Computing (HPC) for RS applications. GDSL is actively collaborating with diverse disciplines including but not limited to Agriculture, Forestry and Natural Resources, Civil Engineering.

EXPERIMENTAL RESULTS (cont'd)

For noisy dataset, we used the training dataset we generated from the cadastral parcel. Pre-trained model using our training dataset was transferred to the other region using hand-digitized clean dataset to verify the transferability. Table 1 demonstrates quantitative results.

Table 1. Quantitative results for our experiment

Region	Precision	Recall	F1-score	IoU (%)
Indiana	0.77	0.71	0.72	64.37
California	0.91	0.90	0.90	82.73

CONCLUSION

We propose a deep learning approach to delineate agricultural parcels using the training data which is generated from pre-existing cadastral parcel. Our results demonstrated the potential possibility of extracting agricultural parcels from pre-existing cadastral parcels. Also, our pre-trained models can be transferred to an area where any parcel information is not established or have extremely low label data, thus facilitating the automation of the time-consuming and labor-intensive digitization process. The proposed approach can be applied to any deep learning models using remote sensing imagery and simple GIS data and can speed up the process of producing the agricultural parcel map.