ASSESSING THE VALUE OF SUPERPIXEL APPROACHES TO DELINEATE AGRICULTURAL PARCELS

MENGGE TIAN
February, 2019

SUPERVISORS:

Prof.Dr.ir. Rolf A. de By
Prof.Dr. Raul Zurita-Milla

ASSESSING THE VALUE OF SUPERPIXEL APPROACHES TO DELINEATE AGRICULTURAL PARCELS

MENGGE TIAN
Enschede, The Netherlands, February, 2019

Thesis submitted to the Faculty of Geo-information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:

Prof.Dr.ir. Rolf A. de By
Prof.Dr. Raul Zurita-Milla

THESIS ASSESSMENT BOARD:

prof.dr. M.J. Kraak (chair)
dr. M.T. Marshall

Disclaimer

This document describes work undertaken as part of a programme of study at the Faculty of Geo-information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Superpixel methods provide special opportunities for delineation of agricultural parcels. For agricultural parcels at smallholder farmland, the situation is more complicated. The project uses multispectral, panchromatic and pan-sharpened images to compare SLIC, SLICO, SNIC, and SNICPOLY. PLS predicted image is also tested to delineate boundary. To achieve better results, aggregation is used as post-process of superpixel generation. This project executes three types of measurements to evaluate the performances of various parameter combinations at pixel level metrics, at area size, and at a shape metric separately.

In general, pan-sharpened images are the best input image in all compared images. However, there is no remarkable diversity among different pan-sharpened images. PLS model does not come out with exceptional results. It reduces the accuracy to some extent. After tests, when the number of superpixels is set around ten times of the ground truth parcel number, the value of compactness is set as $M = 25$, a one-time aggregation delivers the best results.

Keywords

superpixel, smallholder farming, boundary, evaluation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my greatest appreciation to my first supervisor, Dr.ir. Rolf A. de By for his constructive criticisms, contributions and suggestions throughout the whole research phase. He extended his support even during non-working hours and weekends. I am deeply indebted him for firmly supported and guided with language skills and thinking modes.

I sincerely appreciate my second supervisor, Dr. Raul Zurita-Milla for his patience and kind support that helped me a lot in my research methods.

I want to extend my deepest appreciation to my parents for their everlasting love, patience and support spiritually throughout writing this thesis and my my life.

My profound gratitude is deserved by my worldwide friends who support me during the most difficult months.

I am indebted to my amazing friends for their companionship and beautiful memories together.

Thanks to Python for making a hard life sweeter.

Chapter 1
Introduction

1.1 PROBLEM STATEMENT

To achieve the end of hunger by 2030, one of the United Nations' sustainable development goals, "Food security and nutrition and sustainable agriculture" is of great significance (United Nations, 2015). Therefore, it is important to understand agricultural fields, to know where, when and how the crops grow. In developed nations, mature survey techniques and financial support have given sub-cm accuracy in land delineation (European Commission, 2014). In less developed economies, it is infeasible to quantify precisely the magnitude of agricultural parcels. Without such details, policy-makers have a hard time to obtain robust statistics. The ability to delineate the land into farm fields, to identify the crop in each field, to monitor crop health and so forth will be tremendously supportive of building more robust food production systems.

There is the traditional definition of smallholder agriculture (SHA), which is small (less than 2 hectares) and family-run (Lowder, Skoet, & Raney, 2016). Globally, there are more than 570 million smallholder agriculture farms and they cover around 75% of the world's agricultural land (Lowder et al., 2016) and produce 80% of food in less developed countries (Bayer, 2018). Accordingly, supporting SHA farms is vital to provide research-based solutions that sustainably produce food, reduce poverty and end hunger. However, smallholder agriculture is not homogeneous (Fan, Brzeska, Keyzer, & Halsema, 2013), but highly heterogeneous. This means crops grow in highly varied ways (Gollin, 2014), making it hard to identify from images directly. An agricultural parcel is defined as "piece of land larger than 0.25 ha with a minimum width of 30 m" (Food and Agriculture Organization of the United Nations., 2005). The agricultural parcel is proposed as a good unit for SHA research. This project proposes to use agricultural parcels to improve SHA and capacitate to reach their farming potential (Bayer, 2018), to design planning and establish relevant policies (Inan et al., 2010). The delineation of agricultural parcels is assumed to be valuable for diverse agriculture-related agencies. Hence, an elementary requirement is to have delineations at the parcel level.

Thanks to the "Spurring a Transformation for Agriculture through Remote Sensing" (STARS) project, committed to "looking for ways to use remote sensing technology to improve agricultural practices in Sub-Saharan Africa and South Asia" (ITC, 2014), we have available a good collection of WorldView-2 and -3 satellite images providing an image time series. This allows this project to use Very High Resolution (VHR) satellite imagery to extract field boundary information, which may lead to improvements in agricultural practices in sub-Saharan Africa and South Asia.

This proposed research project focuses on segmentation using superpixel generation methods and supervised feature reduction to delineate agricultural parcels from VHR automatically. A superpixel is a kind of image patch as proposed by Ren and Malikn (X. Ren & J. Malik, 2003). The main idea of superpixels is to aggregate a pixel-level area into a district-level area by processing an image with millions of pixels into a few hundred or thousand "superpixels," which is an abstraction of basic image information for an over-segmentation level between pixel level and segmentation level. Another objective of this project is feature reduction. Multispectral image has

eight bands, not all of them are useful or irreplaceable. Feature reduction methods are assumed to reduce redundancy and increase performance for next prediction. In this project, we would like to use it to perform better boundary delineation. Feature reduction has two groups of methods that compress our feature vectors and reduce their dimension. One is feature selection, which is to extract subset variables from multi-variables for delineation parcels to gain more informative than the original image bands. The other group is feature extraction, which applies a transformation on the original feature vector to reduce its dimension.

The aim of the project is to develop an automatic method for agricultural parcel detection to support food security research in areas with smallholder farming systems. The objective is to assess the value of coupling superpixel approaches with supervised feature reduction methods to automatically delineate parcels in very high spatial resolution images. In this project, we want to compare different superpixel approaches, compare approaches with and without feature reduction, compare different evaluation methods and find an optimal approach.

1.2 BACKGROUND

This section presents background knowledge on the creation of superpixels and different kinds of superpixel algorithms.

A superpixel is a group of pixels that has been defined in the domain of image segmentation. Superpixels preserve segment characteristics including texture similarity, brightness similarity, contour energy and good continuation (X. Ren & J. Malik, 2003). Several algorithms exist to generate superpixels, as shown in figure 1.1, named by Achanta (Achanta et al., 2012). State-of-the-art superpixel generation algorithms (Stutz, Hermans, & Leibe, 2018) exist in two categories: graph-based algorithms and gradient-ascent-based approaches.

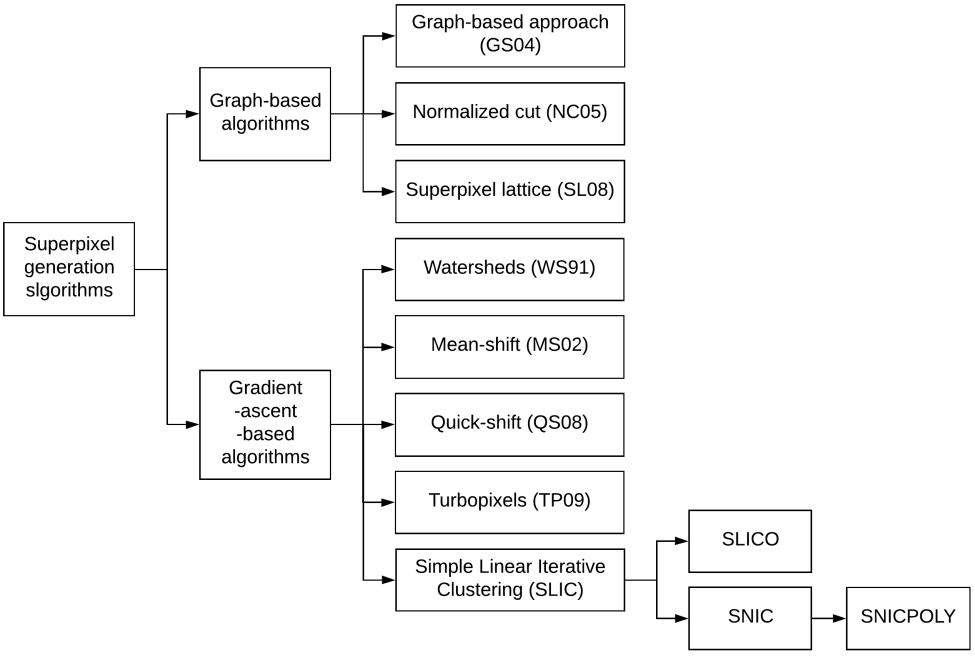


Figure 1.1: Sketch of a superpixel generation algorithms catalogue

Graph-based algorithms recognize the whole image as a weighted undirected graph. Each pixel in the image corresponds with a node in the graph. The adjacency relationship between pixels

corresponds with the edges of the graph. The difference or similarity between pixel features is represented as weight on the edge. Then, the nodes are divided on the basis of various segmentation criteria to complete the segmentation of the whole image. The graph-based approach uses the idea of minimum spanning tree (Felzenszwalb & Huttenlocher, 2004). This algorithm is fast, but cannot control the number of superpixels or their compactness. The normalized cut approach is proposed for measuring the quality of an image partition (Shi & Malik, 2000). This algorithm uses contour and texture features to globally minimize its cost function. It can generate regular superpixels. So, the boundaries cannot be maintained well. Also, it is rather slow for large images. The superpixel lattice approach needs the boundary map of the original image as input, which is reverse with our project objective and cannot be used here (Moore, Prince, Warrell, Mohammed, & Jones, 2008).

For the gradient-ascent-based algorithms, the clustering is coarse at the start, is continuously made finer-grain by the method until convergence is reached. The watersheds algorithm gives a general method and directly generalizes to n-dimensional images (and even to graphs) (Vincent, Vincent, & Soille, 1991). This approach uses topological map to describe an image. The gray value of each pixel (or cost, gradient) of each pixel indicates the altitude of the point.Each local minimum value and its surrounding area are recognized as a basin. The boundary of these basins is formed as superpixels. This approach runs fast and extracts closed contours. But its results can lead to serious over-segmentation of the image. To improve robustness, mean shift is introduced, which analyzes a complex multi-modal feature space (Comaniciu & Meer, 2002), making it run more slowly. Mean-shift approach is based on statistic and iteration. Each point compute with kernel function to get a stable point through limited iteration. Pixels with same stable point are clustered as superpixel (Collins, n.d.). This approach can not set the number of desired superpixels. TurboPixels approach uses geometric flows. Initialized seed points are evenly distributed in the image and dilated based on the change of curvature. For computing a dense over-segmentation of an image, TurboPixels has been proposed and this also contains the boundary of the partial image. Its speed is improved (Levinshtein et al., 2009).

This research project focuses on segmentation detection using the SLIC superpixel algorithm and its extended versions on a larger data frame, as an application in automatic delineation of smallholder agricultural parcels.

1.3 RESEARCH IDENTIFICATION

This research project is about agricultural parcel delineation, and aims to investigate ways to automatically delineate agricultural fields. These agricultural fields contains legumes, maize, millet, rice, sorghum, soybean and so on.

1.3.1 Research objective

The objective is to assess the usefulness of superpixel approaches to automatically delineate parcels in VHR images. In the end, such an automatic superpixel method will be a useful outcome for study areas in sub-Saharan Africa and South Asia. The main objective is divided into the following sub-objectives:

* Delineate agricultural parcels using superpixel approaches.

* Improve with supervised feature reduction

* Evaluate the results

1.3.2 Research questions

To this research objective, the following research questions are answered in this project:

Sub-objective 1 Which image format works best for this project? Which superpixel algorithm has the best performance delineation for SHA?

Sub-objective 2 How does supervised feature reduction work with superpixels? Does supervised feature reduction improve delineation of agricultural parcels?

Sub-objective 3 How can one evaluate the results of agricultural boundary delineation in SHA?

1.3.3 Innovation aimed at

In this project, we use superpixels instead of pixels to delineate agricultural parcels and compare different methods for superpixel generation to find proper parameters automatically at a large scale. There have been several SLIC implementations (scikit, gdal, grass), but seldomly these have targeted agricultural parcel delineation and delineation for SHA. This is considered a more challenging endeavor. With this project, we want to determine whether superpixel algorithms are suitable for agricultural boundary delineation and learn about appropriate values for algorithm parameters in automatic parcel delineation in SHA. Another important innovation is the change from a classical RGB color space, and these superpixel methods working in a new supervised space created using Partial least squares algorithm, which is used as "a projection to latent structures (Wold, Sjöström, & Eriksson, 2001)". This should positively affect the identification of agricultural parcel in complex landscapes.

1.3.4 Method adopted

The workflow in figure 1.2 shows an overview of the method followed to reach the objectives and answer research questions. More details is explained in Chapter 4.

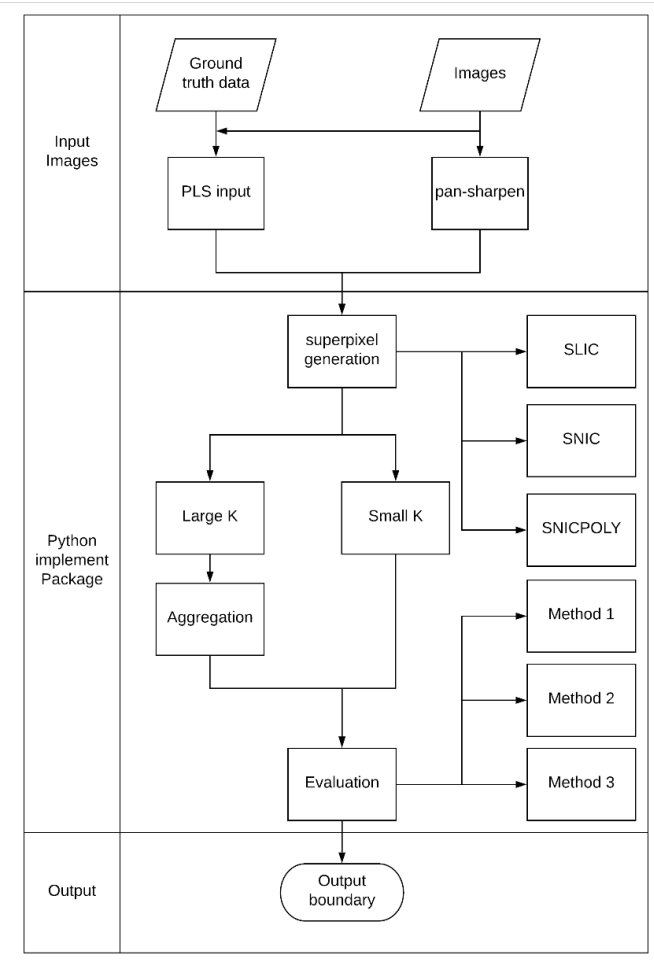


Figure 1.2: The workflow adopted for this research project

1.4 THESIS OUTLINE

Chapter 2 is a literature review on agricultural boundary delineation, superpixel application, feature reduction using partial least squares model and SHA parcel delineation.

Chapter 3 describes the study area, satellite images and reference data, software and hardware used in this research project.

Chapter 4 explains superpixel methods, pan-sharpening process, partial least squares algorithm, aggregation method and evaluation measurements applied in this research.

Chapter 5 presents results with various values of different parameters and evaluation scores of them.

Chapter 6 depicts the general routine of this project and describes general routine of this project and discuss about two issues which are worth continuing to study.

Chapter 7 answers the former questions and summarize conclusions and give some recommendations for further work.

Chapter 2
Literature Review

This chapter provides background information of relevant scientific literature to current state of agricultural parcel delineation, the use of superpixel approaches, feature reduction for object recognition and related work of segmentation.

2.1 AGRICULTURAL PARCEL DELINEATION

For normal agricultural census, land data are collected on parcel level and the total area of each farm owner is derived by summing the areas of all parcels (FAO, 2010). This report makes a distinction between a parcel, a field and a plot and states, which helps this project understand the standard of parcel and common situation. There are paths, cadastral boundaries and/or hedges between parcel by parcel. The ideal situation is that one specific crop is cultivated in one parcel, while mixed cultivation happens. Information on agricultural parcel was pointed out that plays an essential role. Methods of agricultural parcel delineation could be generally divided into three parts: edge detection, line segment detection and object recognition.

Edges in an image have the obvious feature that the brightness or intensity of the image changes sharply (Marr & Hildreth, 1980). Delineation of edge information is a process of detecting the abrupt change of reflection values in an image, and enhancing the differences between the sides. A gradient edge detector has its not-to-be-ignored effect of directionality. The output after application of a gradient filter displays results in some identified direction. Thus, based on normal gradient edge detector techniques, there are sometimes steps as post-processing, like weighted addition filter to improve results (Rydberg & Borgefors, 2001).

Line segments as low-level feature in an image provide core information about the geometric content of the image (Grompone Von Gioi, Jakubowicz, Morel, & Randall, 2010). The ordinary method is to apply an edge detector followed by a Hough transform that extracts all lines that contain a number of edge points exceeding a threshold. The Hough transform is a feature extraction method for detecting curves. It exploits the duality between points on a curve and parameters of that curve. It can be used for arbitrarily complex shape delineation (Ballard, 1981). This transform method gives this project some thoughts to consider.

When considering agricultural parcels as the objects of interest, object-based image analysis (OBIA) can be superior in the identification of complete and closed contours. The challenge that object recognition meets is that an object can cast an infinite number of different 2D images on the canvas, based on the object's position, pose, lightness, background et al. As part of object recognition (Star, 2010), there is a single direction in agricultural parcel delineation: image segmentation analysis (Blaschke, 2010). Segmentation is the "first essential and important step of low-level vision" (Nikhil & Sankar, 1993). Segmentation is a process of partitioning the image into homogeneous regions, and as such is popular in computer vision. It is addressed with various techniques, like the integration of spectral and spatial information (Benediktsson & Ghamisi, 2015).

2.2 SUPERPIXEL APPROACHES

Because of over-segmentation issues in raw imagery, superpixels group pixels similar in color, texture (Reso, Jachalsky, Rosenhahn, & Ostermann, 2013) and other low-level properties. Thus, it captures redundancy in the image, reduces image complexity, and avoids under-segmentation (X. Ren & J. Malik, 2003) in the hope to create visually meaningful regions (Stutz et al., 2018). It also greatly improves the speed of computation. Compared to pixel-based operations, superpixel techniques are good at describing the boundaries between objects and improve speed as a pre-processing step. Treated as a group of connected, similar pixels makes superpixels less sensitive to noise and segmentation scale (Zhang, Jia, & Hu, 2015). Also, superpixels are more flexible and adaptive to local structures of objects in image compared with a regular pixel-based graph (Zhang et al., 2015).

2.3 FEATURE REDUCTION USING PARTIAL LEAST SQUARES

Feature reduction can be categorized as supervised, semi-supervised and unsupervised operations over the input data. Supervised feature reduction methods use labeled data for feature reduction and evaluate feature relevance by measuring the feature correlation with the class label (Sheikhpour, Sarram, Gharaghani, & Chahooki, 2017). Unsupervised methods use data without labels. Thus, these algorithms do not use label information, and this may typically lead to lower performance (Zhao & Liu, 2007). Therefore, these authors propose a method that uses both (a small set of) labeled and (a large set of) unlabeled data in feature analysis to provide "an effective way to address the small-labeled-sample problem."

Feature reduction can be categorized as feature selection and feature extraction. The main idea in these methods is to extract useful and neat data, eliminate irrelevant and redundant data (like noise and duplicates) and improve the characteristic performance of data (Huang, 2015; Chandrashekar & Sahin, 2014; Khalid, Khalil, & Nasreen, 2014).

As one specific feature reduction method, partial least squares (PLS) is a statistical method, which is also called as a structural equation modeling (Haenlein & Kaplan, 2004). The basic PLS algorithm derives a linear model. This model allows two matrices X and Y to project into two different new spaces and builds linear relationship between two new spaces. PLS can handle large data matrices and generates reliable models (Geladi & Kowalski, 1986). Among supervised feature selection methods, ranking by regression coefficients is one of the commonest and simplest ways to select features (Nagaraja & Abd-Almageed, 2015).

2.4 RELATED WORK

Several solid results have been achieved by previous researchers on the development of agricultural parcel delineation, for instance, on how do delineate differently sized objects.

High-resolution remote-sensing images provide substantial and useful information to delineate parcels. In the work of Garcia-Pedrero et al. (García-Pedrero, Gonzalo-Martín, & Lillo-Saavedra, 2017), after superpixel generation, superpixels are used for supervised classification to determine which adjacent superpixels can be merged. The authors use machine learning ideas to achieve the post-processing step, which motivated this research.

Lu et al. (Lu, Oki, Shimizu, & Omasa, 2007) compare several feature extraction and classification methods to map agricultural land using airborne hyperspectral data. The methods they compare include principal components analysis (PCA), decision boundary feature extraction (DBFE), maximum likelihood classification (MLC), extraction and classification of homogeneous objects (ECHO), and image separation into vegetated and non-vegetated area by NDVI. Their work shows

the importance of proper selection of feature methods and the ability of high-resolution images to characterize complex agricultural lands.

Among SHA, the size differences between parcels can be distinctive, setting a condition that superpixel approaches may have a hard time handling. Arbelaez et al. (Arbeláez, Pont-Tuset, Barron, Marques, & Malik, 2014) proposed a method to produce contours with a multiscale technique. First, a high-performance hierarchical segmenter was created to effectively uses multiscale information. Then, they explored efficiently their combinatorial space, combined multiscale regions into highly accurate object candidates and finally, grouped them. The results showed the potential of multiscale combinatorial grouping.

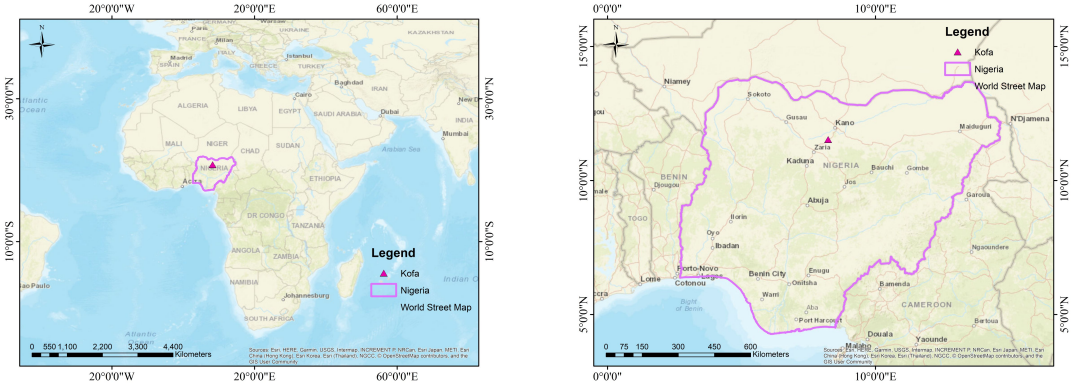
There is another method for differently sized objects using a pyramid attention network (Li, Xiong, An, & Wang, 2018). In this work, spatial pyramid pooling is applied to exploit the multiscale context information for a good descriptor of overall scene interpretation. These models have, however, shown high quality boundaries on several tests while usually needing substantial computing resources and scale choices.

Chapter 3

Data

3.1 STUDY AREA

The study area of this project is a farmland area around Kofa in Nigeria, which is shown in figure 3.1. Kofa is a small village in the Southern Guinea savannah region with a tropical hinterland climate (Ogungbile, Tabo, Duivenbooden, & Debrah, 2012). This area is used to acquire sample images for this research as shown in figure 3.2. For the group of images for this project, the location is between 11°32′ N and 11°36′ N, 8°13′ E and 8°19′ E. It covers an area of approximately 100 square kilometers. The landscape in Kofa is heterogeneous. Household farming is common in both villages and owns high agricultural production potential (Dossou-Yovo et al., 2018). Farming is the main landuse here. The important crop types are legumes, maize, millet, rice, sorghum, soybean, spice crops and vegetables. Mixed crop parcels are common.



(a) Kofa, Nigeria in the world

(b) Kofa in Nigeria

Figure 3.1: Study area

3.2 SATELLITE IMAGES

Due to the common occurrence of parcels less than 2 ha in size, very high spatial resolution images are required. In this projects, WorldView−3 satellite images are used. The geographic coordinate system of all images is WGS1984. The projection coordinate system of all images is UTM, Zone $32N$(EPSG : 32632). The technical parameters of this satellite are shown in table 3.1.

Table 3.1: Technical specifications of satellite images ingested

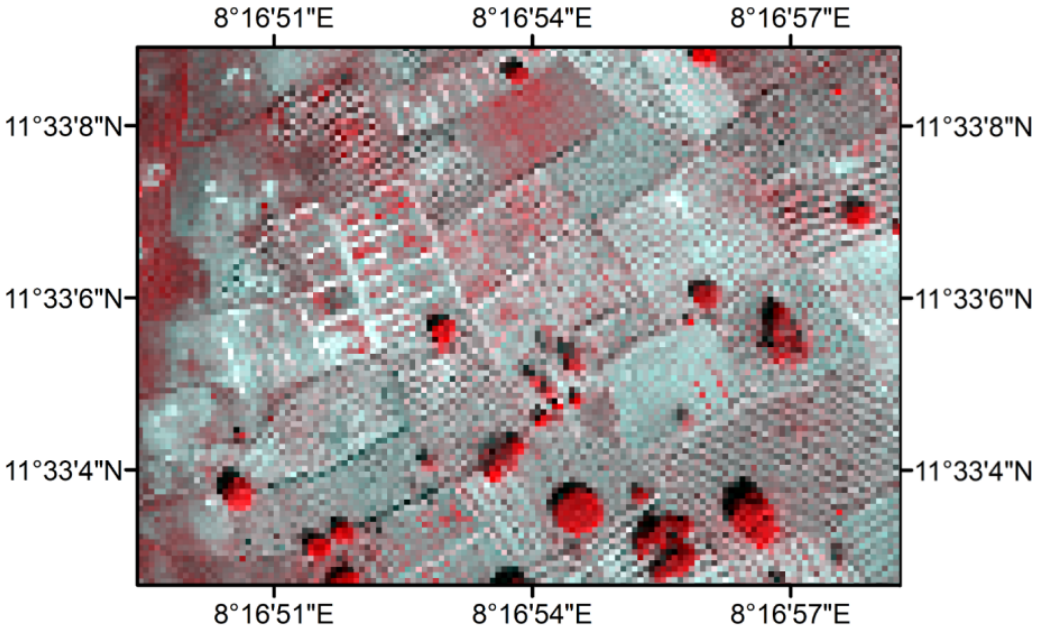
Attribute	WorldView-3
Provider	Digital Globe
Dynamic range	11 bits (14 bits for SWIR)
Panchromatic resolution	0.31 m
Multispectral bands	8 (+8 SWIR)
Multispectral resolution	1.24 m (3.7 m)

Tabl 3.2 shows detailed information of WorldView-3 bands. For the functional test, in the end, images acquired at the same time at different places are required, and images acquired at different times in the same place are needed.

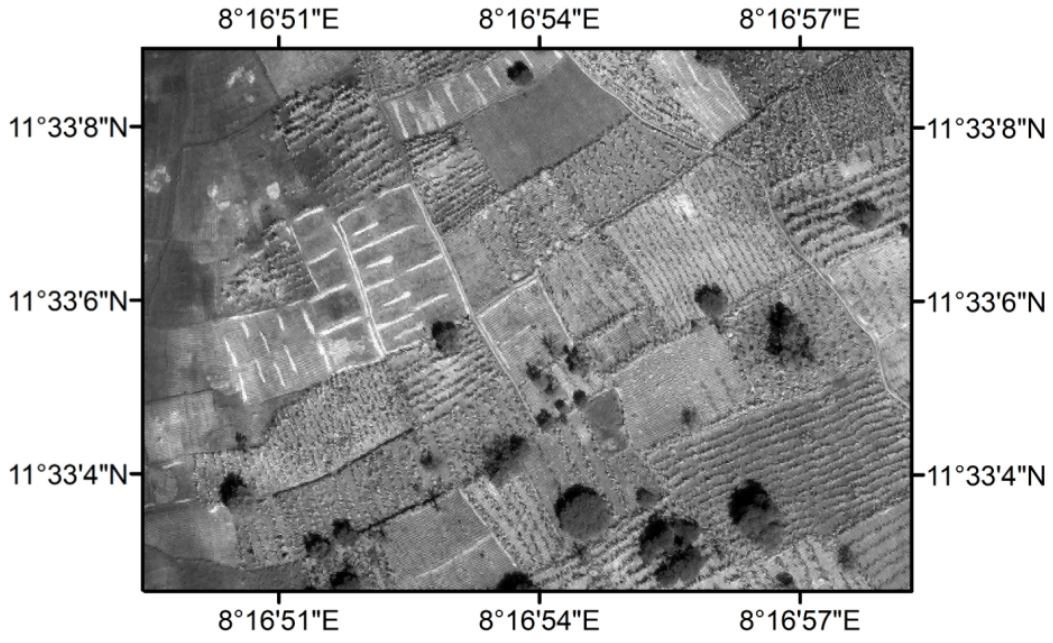
Table 3.2: Band information of WorldView-3

Band no.	Multispectral band	Band wavelengths
1	Coastal	400–450 nm
2	Blue	450–510 nm
3	Green	510–580 nm
4	Yellow	585–625 nm
5	Red	630–690 nm
6	Red Edge	705–745 nm
7	Near-IR1	770–895 nm
8	Near-IR2	860–1040 nm

Our central image data was acquired on November 15th, 2015. It is typically the start of dry season. This is main period of soybean planting, while the planting of maize has been completed. Satellite images used in this project include multispectral images and panchromatic images are all transformed and topology corrected. A sample area is fixed for detailed visualization as shown in figure 3.2.



(a) Multispectral image



(b) Panchromatic image

Figure 3.2: Sample of Satellite images (larger images for comparison with figures in Chapter 5)

3.3 REFERENCE DATA

The reference data is ground truth data covering a part of the Kofa area for 2015. The geographic coordinate system of ground truth shape file is WGS1984. It contains labels of crop system, main crop, landcover and et al. Label of main crop is divided into eight types. Although main crop is labeled, the label of crop system includes "cereal-root crop mixed", "cereal-root crop mixed, unknown main crop", "root an tuber crop" and "undetermined". This mixed plantation would give further parcel delineation a hard time.

The reference data is used for this research as following items:

a) supply training samples for supervised feature reduction.

b) compare with the delineation of agricultural parcels for evaluation.

3.4 SOFTWARE

In this project, the following software is used:

ArcGIS is used for visualization of the position of study area.

Python is used to run superpixel generation methods, feature reduction, and evaluation.

Moreover, the following Python packages would be used:

* several basic packages, like numpy, scipy, matlibplot.

  These packages are used for data organization, math calculation, plot figures and so on.

* several basic geographic packages, like gdal, ogr, rasterio.

  These packages are used for read and write raster and vector images, transform coordinate systems.

* scikit-learn

  This package includes the implementation of SLIC, PLS model.

* shapely

  This package allows spatial analysis, including buffer, intersection, union, calculation of geometry area, bounds and so on.

* geojson

  This package is used to extract and store geometry features, transform data of geometries using Json format and connect shapely package with basic packages.

3.5 HARDWARE

In this research, all of time complexity tests are based on the machine with parameters in table 3.3.

Table 3.3: Basic information of hardware

Item	Description
Processor	Inter(R) Core(TM) $i7-7700$HQ @$2.80GHz$
RAM	16.0 GB
Operating System	Windows 10
System type	64-bit Operating system, $x64$-based processor

Chapter 4
Method

This chapter describes methods that have been used in sequence to achiever the research objectives and address the research questions.

4.1 SIMPLE LINEAR ITERATIVE CLUSTERING

Simple linear iterative clustering (SLIC) is a gradient-ascent-based superpixel algorithm (Achanta et al., 2012) by clustering pixels based on the similarity of color and closeness in xy plane. The image is divided into K segments with the same size S as much as possible. The initial cluster centers are set at the center of each segments, which means the interval between two centroids is also S. A gradient-ascent-based method obtains a better segmentation in every iteration until it converges. The SLIC method convert images from RGB color image space to CIELAB color space. Because the maximum possible distance between two colors in CIELAB color space is limited than the range in RGB space. And CIELAB color space is widely considered as "perceptually uniform for small color distance (Safdar, Cui, Kim, & Luo, 2017)". This method combines a CIELAB color vector and an XY spatial coordinates vector to form a 5-D feature vector $[labxy]$ during initialization. The image firstly is divided into regular grid with the desired number. And recognize the centroid of grid as the initialized centroid of segments. Like k-means, 5D vectors are used to display the location of the pixel in 5-dimensional space and to compute the cost distance, which would both contains the color space and the physical space. For each pixel i in the neighborhood of the centroid C_k, the distance in coordinates space is computed through:

$$D_{coords} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \tag{4.1.1}$$

and the distance in color space is:

$$D_{color} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \tag{4.1.2}$$

These two distance are added together with different coefficients to decide this quantity level of distance as:

$$D(i_k) = \frac{M}{S} \times D_{coords} + D_{color} \tag{4.1.3}$$

where S stands for the size of grid interval and M is used to control the compactness of a superpixel. This weighted sum of distance is the judge statement for each pixel around the cluster center to update the label of current pixel. For each cluster centerC_k, compute each pixel i from $2S \times 2S$ neighborhood around it and compare new computed distance with the stored distance, find the small one and change the label of this pixel to k the small one's k. After all neighbors of all cluster center are computed, one iteration is finished. Used current stored distance value of each pixel to compute new cluster centers and do iteration. After iteration, enforce connectivity is processed to use the region growing method to fine the size of each superpixel in the image and merge too small superpixels into surrounding superpixels. The threshold of small and large sizes could be

adjusted in coding and default as 0.5 and 3 respectively, which means that the sizes of superpixel output are larger than half of the supposed segment size $\frac{XY}{K}$ and smaller than 3 times of that.

In practical, this algorithm add Gaussian filtering before color space conversion. It is found that using it gave better results. And it could be turned off by passing $sigma = 0$.

Algorithm 1: SLIC segmentation algorithm

Input:
RGB image: $I[X, Y]$;
Number of desired segments: K
Output:
label of each pixel in the image: $label[i], i \, \epsilon \, [1, K]$
1. Gaussian filtering of image with adjustable $sigma$ of standard deviation for Gaussian kernel
2. Convert RGB image into CIELAB space $L[l, a, b]$
3. Initialize cluster centers $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ by sampling pixels at regular grid steps S
4. Initialize stored distance attribute of each pixel i: $D_s(i) = \infty$
5. **repeat**
6. **for** each cluster center C_k **do**
7. **for** each pixel i from a $2S \times 2S$ neighborhood around C_k **do**
8. Compute the distance $D(i_k)$ between i and C_k
9. **if** stored distance $D_s(i) > D(i_k)$ **then**
10. Change the label of this pixel to current label $label[i] = k$
11. Update the stored distance $D_s(i) = D(i_k)$
12. Based on $D_s(i)$, compute locations of new cluster centers $C_k(k \, \epsilon \, [1, K])$
13. **until** no pixel changes its distance or maximum number of iterations e has been reached
/*Enforce connectivity*/
14. Set acceptable maximum and minimum scalar of one segment
15. merge superpixels less than the minimum size of segment
16. update the label field with connected values starting at $label = 1$

Specific implementation steps are shown in algorithm 1.

The essential parameter k is needed and indicates the desired number of superpixels. This parameter controls the size of superpixels, and this indirectly affects the shape of superpixels. Other parameters are used internally. The parameter M controls the compactness, and balances color and space distances. A higher M value gives more weight to space distance, so the distance changes largely depend on space distance and makes superpixel shapes more square/cubic, which leads to regular placement and smooth boundaries. This parameter decide strongly image contrast and on the shapes of objects in the image.

All these are targets for fine-tuning in this proposed work. SLIC improves regularity and compactness in the superpixel shapes, and seamlessly accommodates gray-scale as well as color images. Compared with other algorithms, SLIC avoids thousands of redundant distance calculations, which prominently improves the speed (time complexity from $O(kne)$ to $O(n)$, where k is the number of clusters, n is the number of pixels in the image, e is the number of iterations towards convergence). SLIC is also better than other algorithms in compactness, consistency, robustness, depth information and gives superior performance in terms of boundary recall (Stutz et al., 2018).

An extended version of SLIC is SLICO (Achanta et al., 2012); it can be seen as a parameter-free version of SLIC algorithm. SLICO is roughly the same as SLIC except the way M is treated. The user no longer has to set the parameter M of compactness or test different values of it. SLICO adaptively chooses the value of the compactness parameter for each superpixel differently. During

processing, SNICO changes the distance to color in the following way:

$$D_{color} = \frac{\sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2}}{\max(D_{color})} \tag{4.1.4}$$

where $\max(D_{color})$ starts as 1, and compares with D_{color} at each iteration and updates to the larger one immediately after its calculation. At the same time, SLICO keeps the step size S constant. So the distance from a pixel i to the cluster center C_k is calculated with the new distance of color space:

$$D(i_k) = \frac{1}{S} \times D_{coords} + D_{color} \tag{4.1.5}$$

4.2 SIMPLE NON-ITERATIVE CLUSTERING

Another improved version of SLIC is named Simple Non-Iterative Clustering (SNIC) (Achanta & Süsstrunk, 2017). As the name indicates, SNIC does not need iteration. Its approach enforces connectivity of pixels from the start, which means this algorithm requires less memory, runs faster and yet is simpler. The specific implementation is shown in algorithm 2.

Algorithm 2: SNIC segmentation algorithm

Input:
RGB image: $I[x, y]$
Number of desired segments: K
Output:
label of each pixel in the image: $label(i), i \, \epsilon \, [1, K]$;
1. Initialize $label(i) \leftarrow 0$
2. Convert RGB image into CIELAB color space $L[l, a, b]$
3. Initialize cluster centers $C_k = [l_k, a_k, b_k, x_k, y_k]^T$
4. **for** $k \, \epsilon \, [1, K]$ **do**
5. Initialize eltement $e \leftarrow \{C_k, k, 0\}$
6. Build a priority queue Q, push e on th equeue Q
7. **while** Q is not empty **do**
8. Pop Q to get e_i
9. **if** $label[i]$ is 0 **then**
10. $label[i] = k_i$
11. Update centroid $C[k_i]$ online with C_i
12. **for** Each connected neighbor j of i **do**
13. **if** $label[i]$ is 0 **then**
14. Compute the distance $d_{j,k_i} = \sqrt{\frac{\|x_j - x_k\|_2^2}{s} + \frac{\|c_j - c_k\|_2^2}{m}}$
15. Create element $e_j = \{x_j, c_j, k_i, d_{j,k_i}\}$
16. Push e_j on Q
17. **return** L

Instead of using iteration, this method takes advantage of a FIFO queue data structure. The queue used is a lists with first-in-first-out (FIFO) behavior. It is based on the principle that only inserts happen only at the back (the rear terminal position, known as *enqueuer*) and deletes only happen at the front (the front terminal position, known as *dequeuer*). When queue Q is not empty, the top-most element is popped to operate.

SNIC runs in a single iteration, does not use distance map and requires less memory. The Berkeley 300 dataset is a valuable resource for studying statistics of natural images (Martin, Fowlkes, Tal, & Malik, 2001). SNIC tests done by Achanta and Süsstrunk (Achanta & Süsstrunk, 2017) on the Berkeley 300 dataset showed that this algorithm performs better than state-of-the-art algorithms, including SLIC.

4.2.1 The Douglas-Peucker algorithm

An extended version of SNIC is named SNIC-based polygonal partitioning (SNICPOLY) (Achanta & Süsstrunk, 2017). This method uses the Douglas-Peucker algorithm (Douglas & Peucker, 1973) during post-processing to smoother boundaries. The Douglas-Peucker algorithm simplifies curves, which helps to detach from pixel boundaries, leads to fewer boundary vertices and uses less calculation and memory.

An example of the DP algorithm is illustrated in figure 4.1 (Mokrzycki & M, 2012). In this sketch, the vectices n_1 and n_p are the first and the last points on a curve. In the first step, the straight line from n_1 and n_p is drawn. Then the longest perpendicular distance (x) to the straight line n_1n_p is calculated from the vertex n_k. If x sits above the defined threshold value ϵ, the straight line n_1n_p is separated into two straight lines n_1n_k and n_kn_p. If not, the vertex is skipped. Thus, after going over every vertex in the curve, the line with the vertices $n_1 \cdots n_p$ is reduced to a number of line segments (figure 4.1(e)). The defined threshold, also known as epsilon, is a vital parameter of Douglas-Peucker Algorithm, and we call it the Douglas-Peucker number (dpn) in this project.

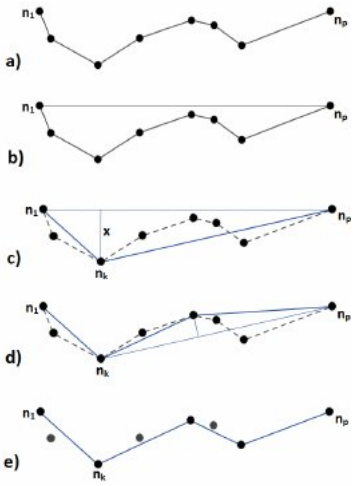


Figure 4.1: Sketch of the Douglas-Peucker algorithm

Due to pixel-level clustering, superpixels are generated with no doubt a lot of small oscillation at the edges. After processing Douglas-Peucker algorithm, the boundaries become gentle. Moreover, in this case, agricultural parcel always have regular edges similar to rectangles. When dpn is set as large as the short side of parcel, this algorithm gives a better simulation of agricultural boundaries. Therefore, the Douglas-Peucker algorithm obtains good results especially as post-process after the above superpixel algorithms.

4.3 THE PAN-SHARPENING PROCESS

In the study area of this research project, raw data is available in the form of a WorldView-3 satellite multispectral image and a panchromatic image. The multispectral image has a resolution of

1.24 m at nadir, and the panchromatic image has 0.31 m at nadir. Pan-sharpening is shorthand for panchromatic sharpening. It means using a high-resolution panchromatic (single band) image and a low-resolution multispectral image to derive a sharpened multispectral image. The "sharpen" means to increase the spatial resolution of the multispectral image by that of the panchromatic image (Padwick, Deskevich, Pacifici, & Smallwood, 2010).

One common class of algorithms for pan-sharpening is called "component substitution," and this include general pan-sharpening methods for WorldView images, like Brovey, Esri, IHS, Simple mean transformation and et al (Padwick et al., 2010).

In the following equations, P stands for panchromatic image, B, G, R for red, green and blue band of input image, respectively, NI for near infrared band of input image and W_i for required weight of band i.

In the Brovey transformation, the general equation uses R, G, B and the panchromatic bands (P) as inputs to output new red, green, and blue bands. Take red band as an example as follows in equation 4.3.1. The green and blue band are computed in a similar way:

$$\text{Red}_{out} = \frac{R_{in}}{(B_{in} + G_{in} + R_{in}) \times P_{in}} \tag{4.3.1}$$

The Esri pan-sharpening transformation uses a weighted average W_A to create an adjustment value (ADJ) to compute its pan-sharpened outputs. Take the red band as an example in equation 4.3.2. The green and blue bands are computed in a similar way and the near-infrared band could also be used as an optional result:

$$\begin{aligned} \text{ADJ} &= P - W_A; \\ \text{Red}_{out} &= Red_{in} + \text{ADJ} \end{aligned} \tag{4.3.2}$$

IHS is the abbreviation of intensity-hue-saturation transformation. The intensity band can be computed as follows , using equation 4.3.3:

$$I_{out} = P_{in} - I_{in} \times W_I \tag{4.3.3}$$

The simple mean transformation uses a simple mean average of the input band. This is an example of using the red band in equation 4.3.4:

$$\text{Red}_{out} = 0.5 \times (\text{Red}_{in} + P_{in}) \tag{4.3.4}$$

Another common class of algorithms for pan-sharpening is called "interpolation," including nearest-neighbor, bilinear, bicubic, Lanczos interpolation and so on (Opencv dev team, 2019). Some of these methods are shown in the following figure 4.2:

Nearest-neighbor interpolation finds the closest sample from the subset of input to the query point and applies a piecewise value. This method may cause discontinuity in the gray scale of the image generated by interpolation, and may appear obviously jagged in the place where the gray scale changes.

Bilinear interpolation is a 2-D extension of linear interpolation for bivariate functions on x and y axes. In a $2D$ image, assume i, j are non-negative integers, and u, v are floating point numbers in the range of $[0, 1)$. For the target point $(i + u, j + v)$, the interpolated value $f(i + u, j + v)$ is determined by a $2D$ plane consisting of 4 points around it.

Bicubic interpolation consider 4×4 pixel neighborhood around a floating point $(i+u, j+v)$. The value of target point is computed with a $3D$ surface consisting of 16 points around it.

Lanczos interpolation is a filtering algorithm over an 8×8 pixel neighborhood. This method uses horizontally stretched sinc function to simulates smooth interpolated values. The sketch simply display its result in figure 4.3.

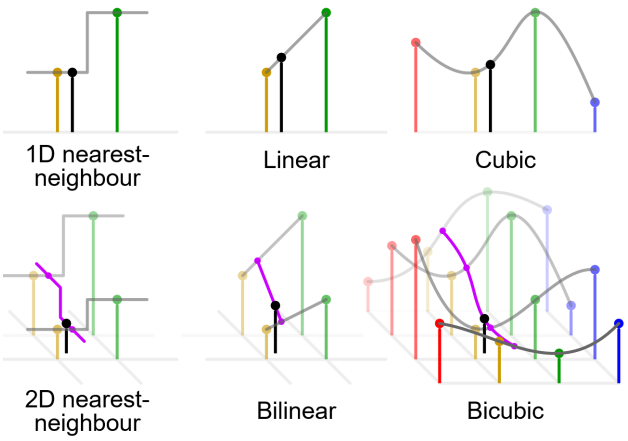


Figure 4.2: Sketch of some interpolation methods (Cmglee,2016)

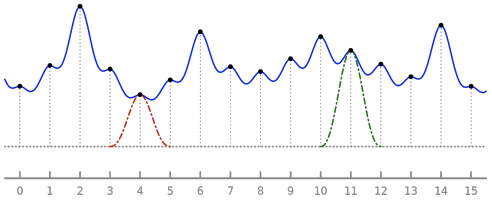


Figure 4.3: Sketch of Lanczos interpolation (black points are original values, solid blue curve is its Lanczos interpolation, Jorge Stolfi, 2013)

4.4 PARTIAL LEAST SQUARES ALGORITHM

As one of the feature reduction techniques, the partial least squares algorithm (PLS) uses the idea of dimensionality reduction. PLS is an extension of multivariate linear regression. Many multivariate linear regression methods, including discriminant analysis (DA), Canonical correlation analysis (CCA), principal components regression (PCR) have two important attributes in common. These methods impose restrictions such that: (1) the basic factors of Y and X variables are extracted from $Y'Y$ and $X'X$ matrices, respectively, rather than from the cross product matrix involving the Y and X variables; (2) the number of prediction functions never exceeds the minimum of the Y variable and the X variable. Partial least squares regression extends multiple linear regression without imposing the restrictions employed by above-mentioned methods. In partial least squares regression, prediction functions are represented by factors extracted from the $Y'XX'Y$ matrix. To be clear, this algorithm considers extracting the principal components of Y and X as much as possible (PCA), and considers extracting X and Y separately and maximizes the correlation between the principal components (CCA).

The general underlying model of multivariate PLS is provided in equation 4.4.1:

$$\begin{aligned} X &= TP^T + E \\ Y &= UQ^T + F \end{aligned} \tag{4.4.1}$$

where X is an $n \times m$ matrix of predictors, Y is an $n \times p$ matrix of responses; T and U are $n \times l$ matrices that are, respectively, projections of X (the X score, component or factor matrix) and projections of Y (the Y scores); P and Q are, respectively, $M \times l$ and $p \times l$ orthogonal loading matrices; and matrices E and F are the error terms, assumed to be independent and identically

distributed random normal variables. The decompositions of X and Y are made so as to maximise the covariance between T and U.

In this research project, a PLS algorithm can use the crop label for some of the delineated parcels as the data of variable Y and multispectral bands as data of variable X.

4.5 AGGREGATION APPROACH

Although SLIC enforces connectivity for initially generated superpixels, that part mainly considers the size of superpixels. This part's aggregation approach focuses on the position relationship and statistical analysis of superpixels as discriminant conditions. This aggregation approach is implemented as a post-processing step following superpixel generation.

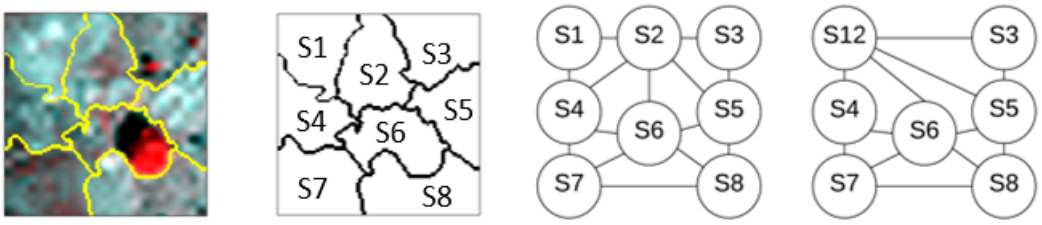


Figure 4.4: Sketch of superpixel merger

The method, as the sketch shows in figure 4.4, through labeled image segments, builds an undirected graph that includes adjacency relationships between pairs of superpixels. Each vertex in the graph stands for a segment. We give each pair of connected segments a edge and compute the cost of paired segments as edge's weight. The cost is calculated though equation 4.5.1.

$$\text{Cost} = (\bar{R}_i - \bar{R}_j)^2 + (\bar{G}_i - \bar{G}_j)^2 + (\bar{B}_i - \bar{B}_j)^2 < T \tag{4.5.1}$$

For a pair of connected segments i, j, the mean values of three bands $\bar{R}$, $\bar{G}$ and $\bar{B}$ are calculated and used to compare. Threshold value is tested and stated to determine whether we should merge this pair of superpixels. This aggregation judgment condition is inspired by "fast region merging" (Haris, Efstratiadis, Maglaveras, & Katsaggelos, 1998). The value of threshold T is selected by tests.

4.6 EVALUATION OF RESULTS

To evaluate the results of superpixel segmentation, this project has identified three methods to assess results at different levels as follows. The first looks at pixel level metrics, the second at area size, and the last at a shape metric.

4.6.1 Boundary recall

At pixel level, this project imply confusion matrix of segment boundary and ground truth boundary. We are setting a desired distance threshold d to define an acceptable buffer size around boundaries. Buffered line sections (to avoid misunderstanding, we name it line section instead of line segmentation) from segmentation S are compared with buffered line section from polygons from ground truth data G. We consider as True Positive (TP) the number of pixels inside the intersection area; True Negative (TN) is the number of pixels that are within the ground truth buffer but not within the segmentation line buffer; False Positive (FP) is the number of pixels not within the

ground truth buffer but in the segmentation line buffer, and False Negative (FN) the number of pixels outside of any line buffer. The boundary recall scores are defined in equations 4.6.1, 4.6.2, and 4.6.3 as follows.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4.6.1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4.6.2}$$

$$\begin{aligned} \frac{2}{\text{F1}} &= \frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}, \\ \text{F1} &= \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \end{aligned} \tag{4.6.3}$$

These three scores always fall between 0 and 1. The higher score, the better the result.

4.6.2 Segmentation accuracy

At area level, the achievable segmentation accuracy(ASA) is chosen to access the accuracy of superpixel segmentation. These scores are determined with equation 4.6.4.

$$\begin{aligned} \text{ASA}_s &= \frac{\sum_i \max_j (S_i \bigcap G_j)}{\sum_i G_i}, \\ \text{ASA}_g &= \frac{\sum_i \max_j (G_i \bigcap S_j)}{\sum_i S_i} \end{aligned} \tag{4.6.4}$$

S_i is a segment created by the algorithm and G_j is a ground truth segment. The subscript i and j do not have to be different. The formula $(S_i \bigcap G_j)$ indicates the area of the intersection. For each segment S_i from created segments, the maximum intersection of a certain ground truth segment G_j is found. ASA_s is the quotient of the sum of all intersected area divided by the sum of all S_i. ASA_s measures the fraction of ground truth area that is correctly represented by superpixel methods, while ASA_g measures the fraction of superpixel area that is correctly labeled by ground truth segments.

These two scores express two qualities of correspondence relationship, while ASA_s is also known as "upper-bound segmentation accuracy." Both have values between 0 and 1. Also, here we have that the higher score, the better the result. As ASA_s grows, ASA_g reduces. This would be hard to determine a better result. An index, the product of ASA_s and ASA_g is considered as the determine condition. the The graph for $\text{ASA}_s - \text{ASA}_g$ is created, using ASA_s as X axis and ASA_g as Y axis. The area of rectangles at the origin coordinates and the point (ASA_s, $\text{ASA}_s - g$) indicates the good or bad of the results. The larger area, the better the result.

4.6.3 Shape and distance metrics

At shape level, to compare the trajectories, much research has been done in the past, like formerly used point-based distance, Euclidean distance. Point-based distance methods also include: dynamic time warping (DTW), longest common subsequence (LCSS), Edit distance on real sequence (EDR). Other than these, there are shape-based distance metrics: Hausdorff distance, Frechet distance; segment-based distances: One Way Distance, LIP distance; and task-specific distances: TRACLUS, road, semantic (Wang & Sadiq, 2013).

In this project, the Frechet distance is used to evaluate the ground truth line and the generated (boundary) line segments. Compared with other methods, the Frechet distance considers in detail

the continuity of shape. This method has an intuitive definition: suppose a man is walking his dog. Assume the man and the dog walk along different curves with different speeds, and they are not allowed to move backwards. The Frechet distance is the minimum length of a leash that the man and the dog ties during their walk. A sketch is shown to illustrate in figure 4.5 (Chen, Ma, Jing, Guo, & Xiong, 2017).

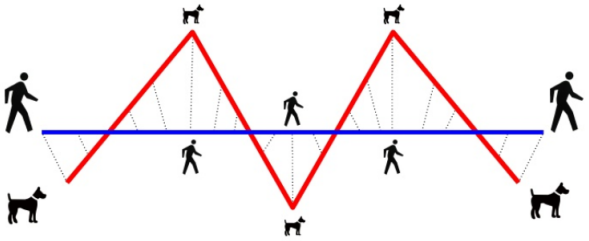


Figure 4.5: Sketch of Frechet distance

In this project, the superpixel list of curves is generated from the closed boundary of a segment, while the ground truth list of curves is created from boundaries from ground truth shapes. Frechet distances are found for each pair of superpixel curve and ground truth curve. The smallest distance is considered as the quantity to compare results from different situations.

Chapter 5
Results

We have fixed a 390×542 image inside the whole image as an example for detailed visualization, which has been described in chapter 3. A combination of bands 7, 5, 2 is used for false-color synthesis to pan-sharpen and extract color information as much as possible the various agriculture types. Before showing all the results, ground truth boundary of sample area is shown in figure 5.1 to compare with following superpixel boundaries.

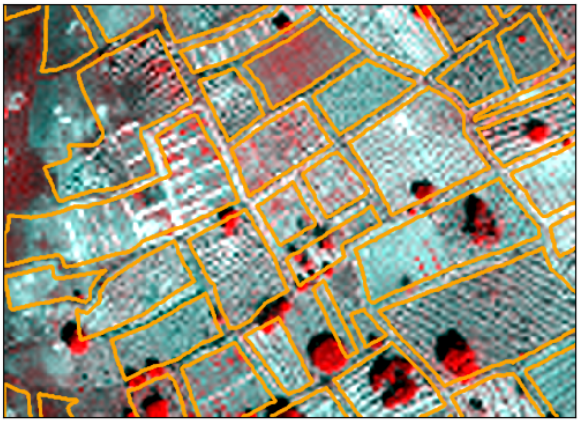
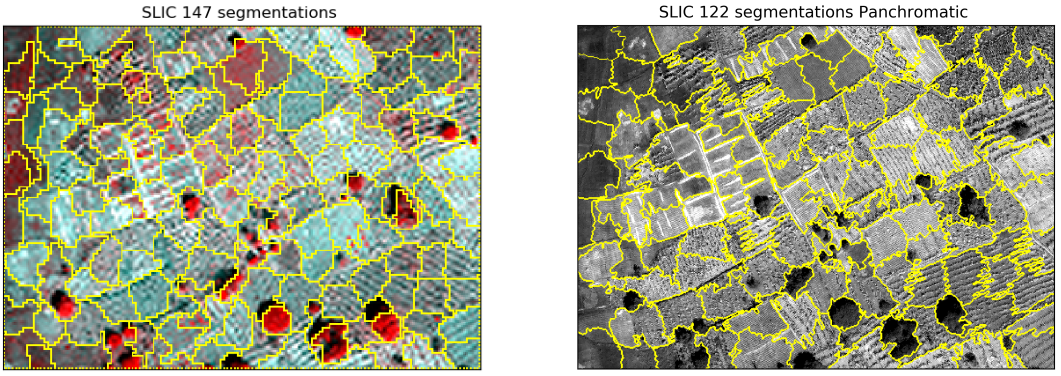


Figure 5.1: Ground truth boundary (orange lines)

The test of boundary generation starts at SLIC algorithm with $K = 150$ and $M = 15$. The results of multispectral and panchromatic sample images are presented in figure 5.2 separately. Yellow lines represents boundaries (All yellow lines in following figures are all generated boundaries).



(a) Boundary from multispectral image ($K = 150, M = 15$)

(b) Boundary from panchromatic image ($K = 150, M = 15$)

Figure 5.2: Multispectral and panchromatic images generated boundaries

5.1 PAN-SHARPENED IMAGES

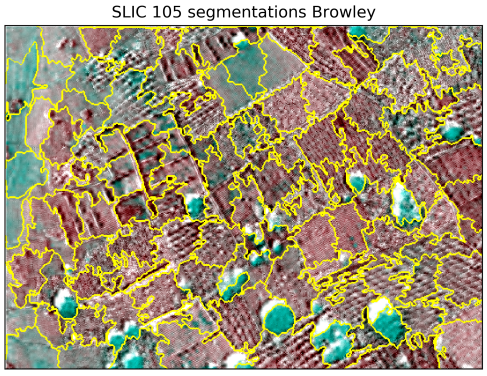
The pan-sharpened results of various methods are shown in figure 5.3 for component substitution methods and in figure 5.4 for interpolation methods.

In this project, four types of component substitution methods are discussed. With over-segmentation, four types pan-sharpened images all return some good boundary delineation. The Brovey method increases the visual contrast of high and low parts in the data histogram. The Esri method uses weighted averaging to adjust the result. By the weight value, the green output is made more vibrant. After IHS component substitution, the sharpness of the image is increased compared with the original multispectral, while the Brovey method increases it less. Brovey and IHS pan-sharpening images have prominent boundary features, and also the texture of different agricultural parcels. This characteristics is especially prominent in the mixed cropland. When generate super-pixel, it is hard to determine whether texture line or boundary line.

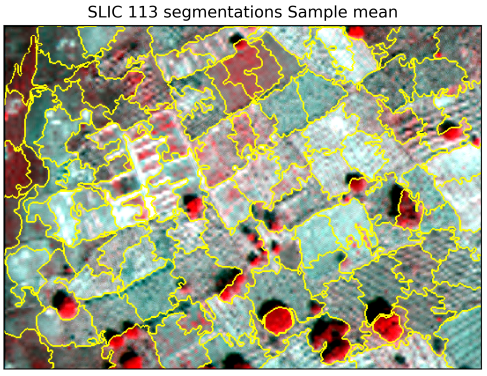
In this project, four types of interpolation methods are discussed. Preferable interpolation methods suggested by an official tutorial (Opencv dev team, 2019) are bilinear and bicubic (slow) for zooming, especially for zoom factors below three. Among four typical interpolation methods tested in this project, nearest-neighbor interpolation is simple in computation and requires little memory, but the edge of the parcels and changes of different agricultural crops, looks more grainy. Bilinear, bicubic and Lanczos methods overcome the shortcoming of disconnected image values. However, the edges obtained by the bilinear method are smoothed and contours may blur. The bicubic method comes with the effect of edge enhancement and better maintains the structure of the image. The Lanczos method also enhances the edge, but displays too much detail of texture in the image.

To give a qualitative comparison, three types of evaluation are used to compare the results with the condition that K is set from 50 to 450 while $M = 15$ in figure 5.5.

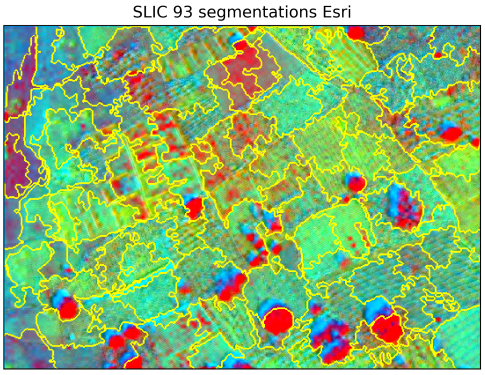
Pan-sharpened images get better results than multispectral image, especially $F1$ scores from figure 5.5(a). Various pan-sharpened methods return similar results, therefore pan-sharpened method choice does not really matter. Based on figure 5.3, brovey result could incorrectly mixed texture area and plain area. Sample mean and Esri results could not describe edges of parcels contain two types of crop. Ihs results could not draw edges of two texture parcels. The results based on bicu-



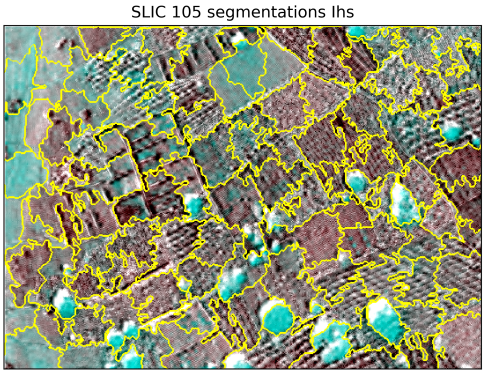
(a) Brovey pan-sharpening ($K = 150, M = 15$)



(b) Sample mean pan-sharpening ($K = 150, M = 15$)



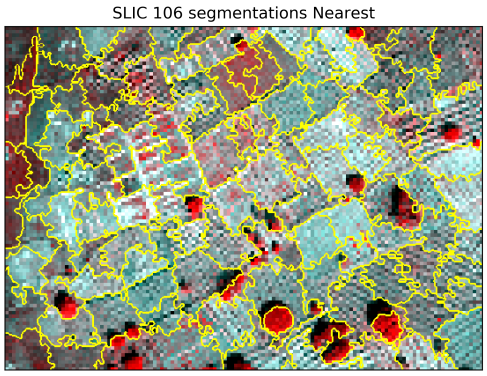
(c) Esri pan-sharpening ($K = 150, M = 15$)



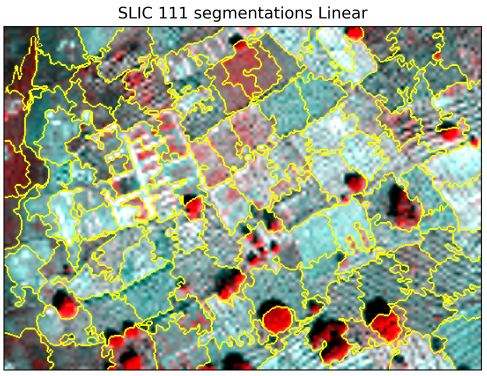
(d) IHS pan-sharpening ($K = 150, M = 15$)

Figure 5.3: Pan-sharpened images obtained from various component substitution methods

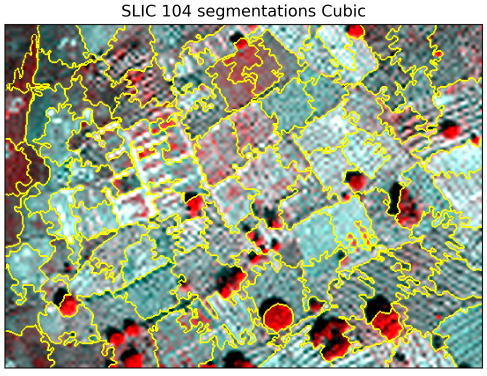
bic method nearly determine all boundaries or have a little shift. Bicubic method is chosen for forward steps for making edges of parcels easier to interpret.



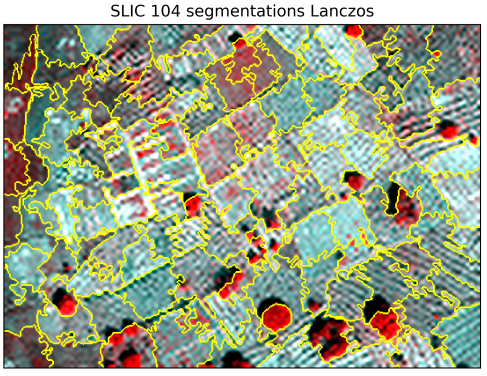
(a) Nearest-neighbor pan-sharpening ($K = 150, M = 15$)



(b) Bilinear pan-sharpening ($K = 150, M = 15$)

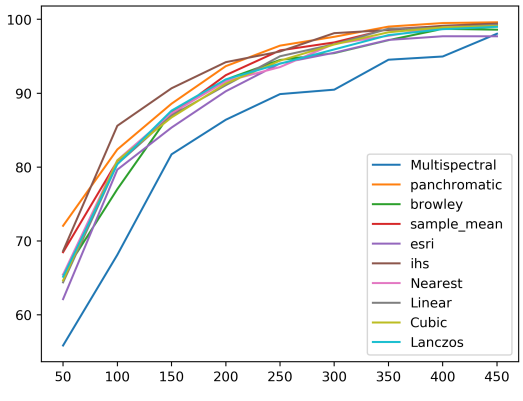


(c) Bicubic pan-sharpening ($K = 150, M = 15$)

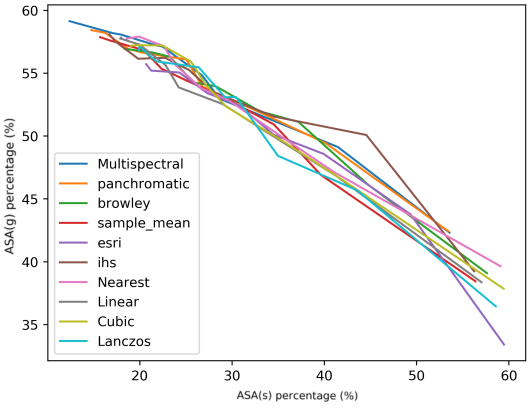


(d) Lanczos pan-sharpening ($K = 150, M = 15$)

Figure 5.4: Pan-sharpened images obtained from various interpolation methods



(a) $F1$ score graph of various methods

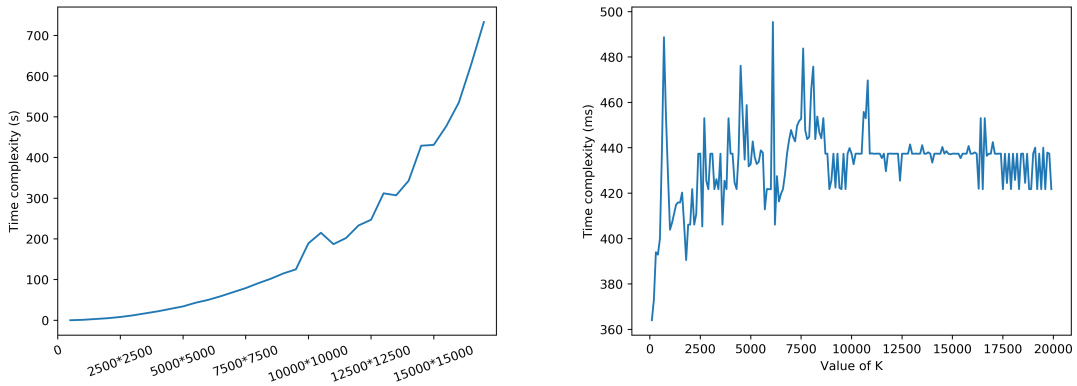


(b) $\mathrm{ASA}_s - \mathrm{ASA}_g$ graph of various methods

Figure 5.5: Evaluation results of various pan-sharpening image products

5.2 TIME COMPLEXITY OF SLIC

Because of the image sizes involved, the time complexity of superpixel generation may become an important feature. To see the time complexity on variously sized $n \times n$ images, the values of K and M are kept fixed at 350 and 15, respectively. The results are presented in figure 5.6. As n grows from 10 to around $10,000$, the complexity continuously increases. When n is larger than $10,000$, the time cost increases with some small oscillations, which happens in repeated tests (figure 5.6(b) just displays one time result). When n is larger than $17,500$, the machine mentioned in the hardware section could not run the process because of memory error.



(a) Time complexity of different size of image

(b) Time complexity of different value of K

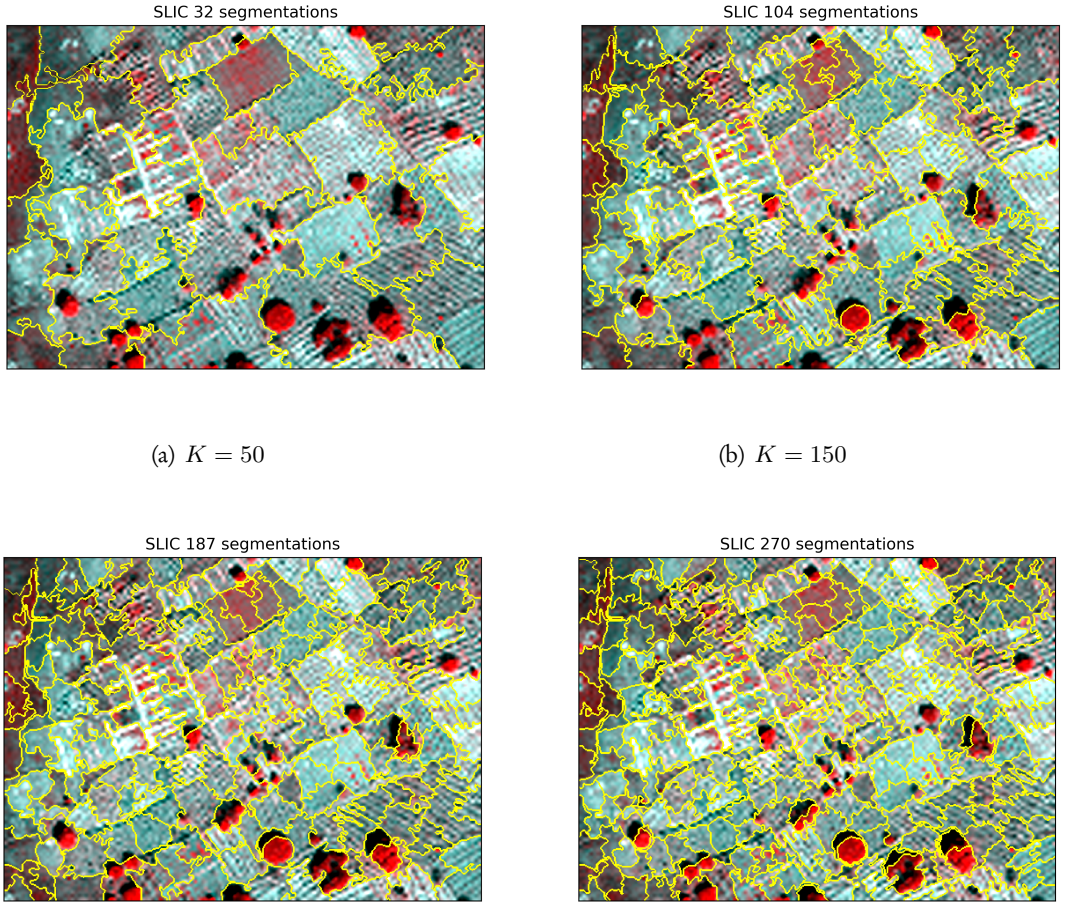
Figure 5.6: Time complexity tests

To see the change in performance with different values of K, the test image is fixed to that of the example used in the pan-sharpening section, and the value of M is kept fixed at 15. The result is shown in figure 5.6. In practice, Cython is used to improve enforceability. Cython is a C-extension for Python. It is written in Python code that calls back C or C++ code natively at any point, which speeds up the execution of Python code. With Cython acceleration, although the time cost have the tendency of increase along the increase of K, it still less than one second in the range researched (K less than $20,000$).

5.3 TUNING OF SLIC'S K PARAMETER

The first essential parameter of SLIC is the number of requested superpixels K. Because the SLIC algorithm enforces connectivity (described in chapter 4) by merging segments that are smaller than set size, the number of resulting segments is always a bit smaller than the requested K. The results with various values for K are shown in figure 5.7 (all yellow lines are superpixel boundaries). As a cluster algorithm, the number of desired cluster center K influences the boundary of the results. It is obvious when comparing these four different K results.

When K is small, in this case 50, some of the boundaries provide a good shape description of some single parcels. But when agriculture is mixed or the area contains trees, the segments may become confused and mixed together or be divided into different segments. This situation is less frequent with increasing K. For instance, when K is 150, though tree shadows are no longer detected as single segments, the shapes of trees come out better. When K is 200, shapes of trees are well-portrayed and so are some of their shadows. In the current situation, in spite of bad visualization and divided parcels, superpixels provide good description of the parcel edges,



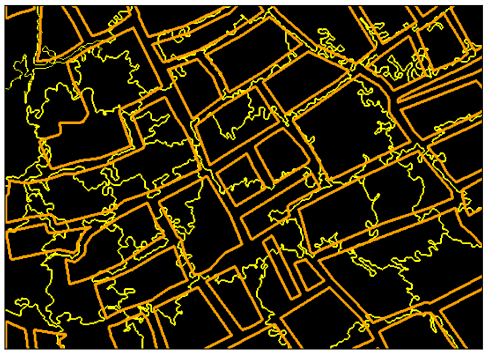
(a) $K = 50$ (b) $K = 150$

(c) $K = 250$ (d) $K = 350$

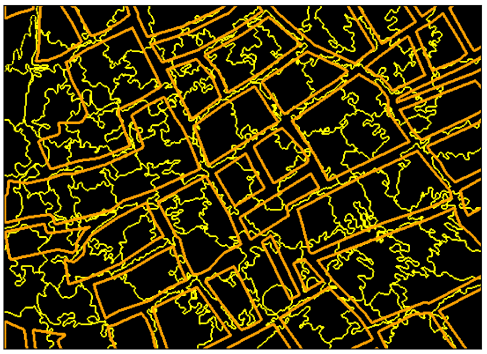
Figure 5.7: Boundaries for different value of K with fixed $M = 15$

which is also a reason to do aggregation later. When K is 250, shapes of tree shadows remain well-depicted. When K is larger, more detail is shown on the boundaries, but also with an increase in useless boundary lines, which provides an opportunity for aggregation methods.

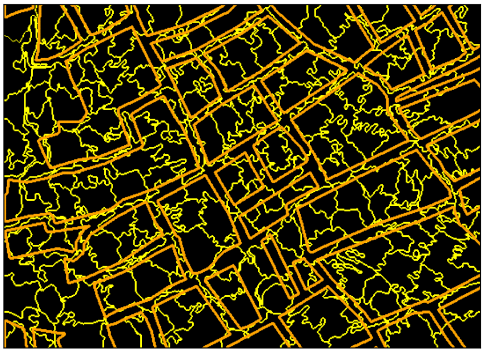
All graphs in figure 5.12 share the legend of (c)$F1$. Overall, except curves with $M = 5$, curves have similar tendency in all evaluations. When $K = 150$ and $M = 10$, the result gets the best precision, recall and $F1$ score. When $K = 350$ and $M = 15$, the result get shortest Frechet distance. And this pair of values also performs well in precision, recall and $F1$ score. Therefore, $K = 350$ and $M = 25$ are chosen as default value for further tests.



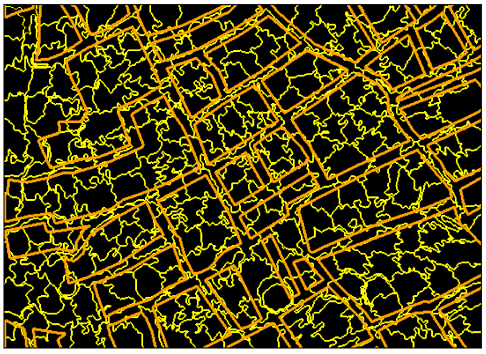
(a) $K = 50$



(b) $K = 150$

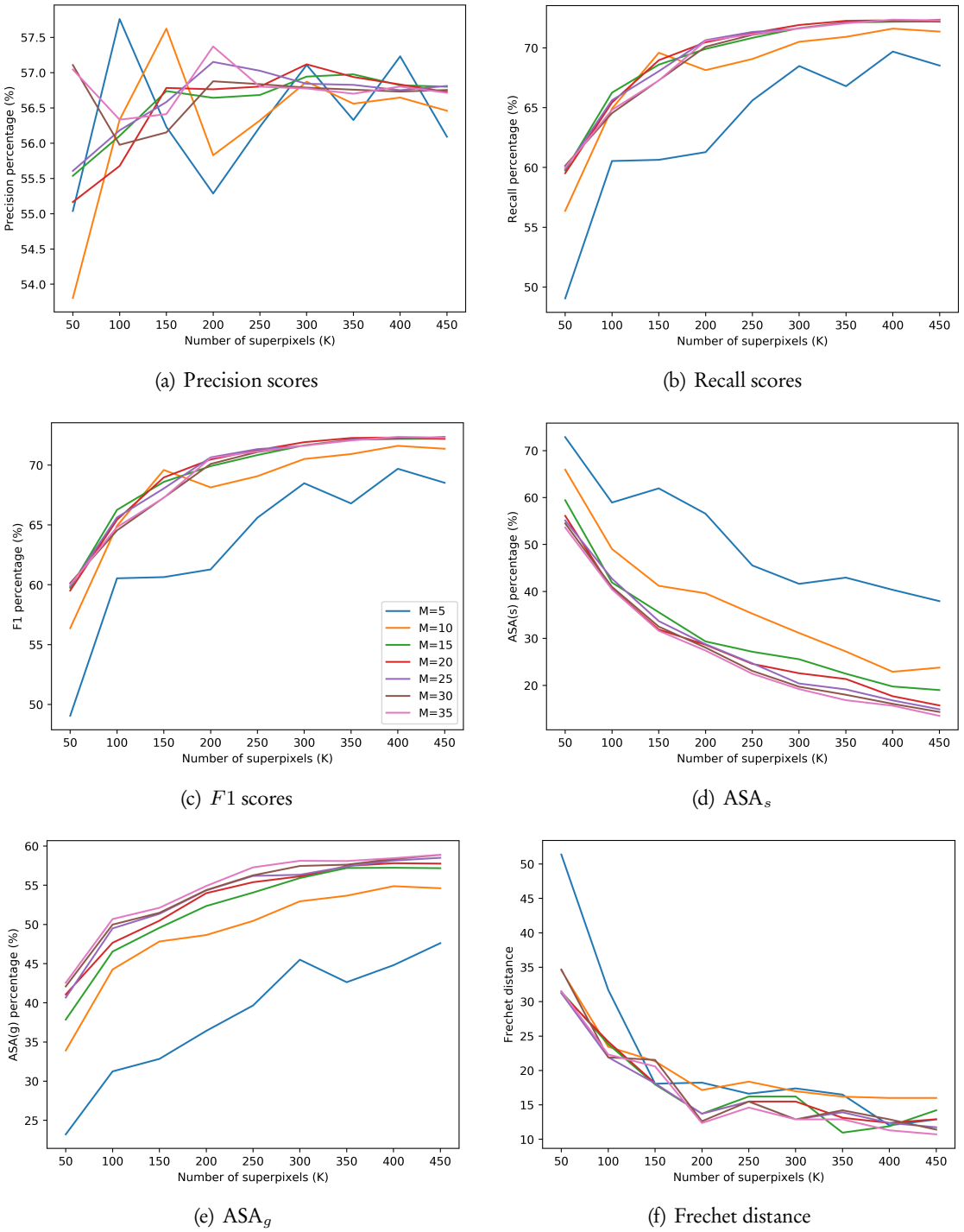


(c) $K = 250$



(d) $K = 350$

Figure 5.8: Boundary comparison with ground truth parcels for different values of K with fixed $M = 15$ (yellow lines are segment boundary, orange lines are ground truth boundary)



(a) Precision scores

(b) Recall scores

(c) $F1$ scores

(d) ASA_s

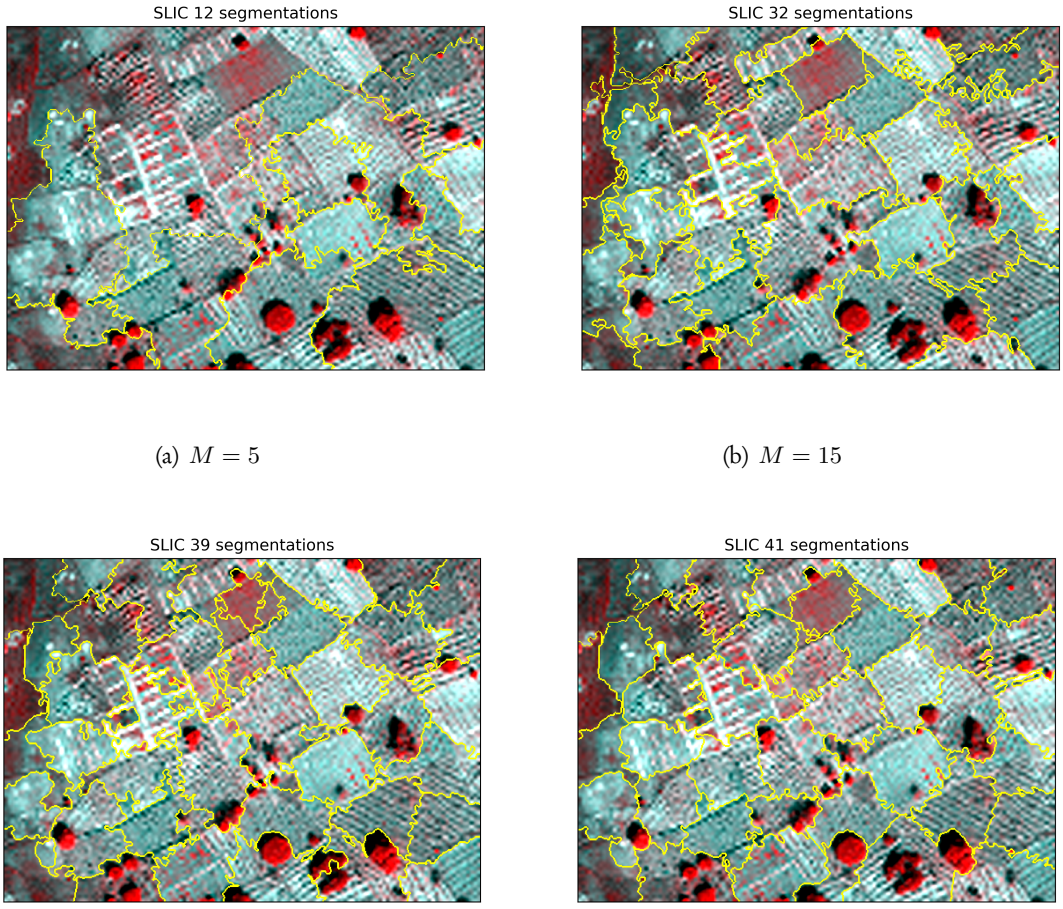
(e) ASA_g

(f) Frechet distance

Figure 5.9: Evaluation graphs for different value of K with fixed M

5.4 TUNING OF SLIC'S M PARAMETER

This section studies the value of compactness M, as introduced in equation 4.1.3. This parameter allows to tune the balance between distance in coordinate space and distance in color space. The results, using various values of M with fixed number of superpixels $K = 50$ in figure 5.10 and fixed $K = 350$ in figure 5.11, which shows the influence of this parameter.



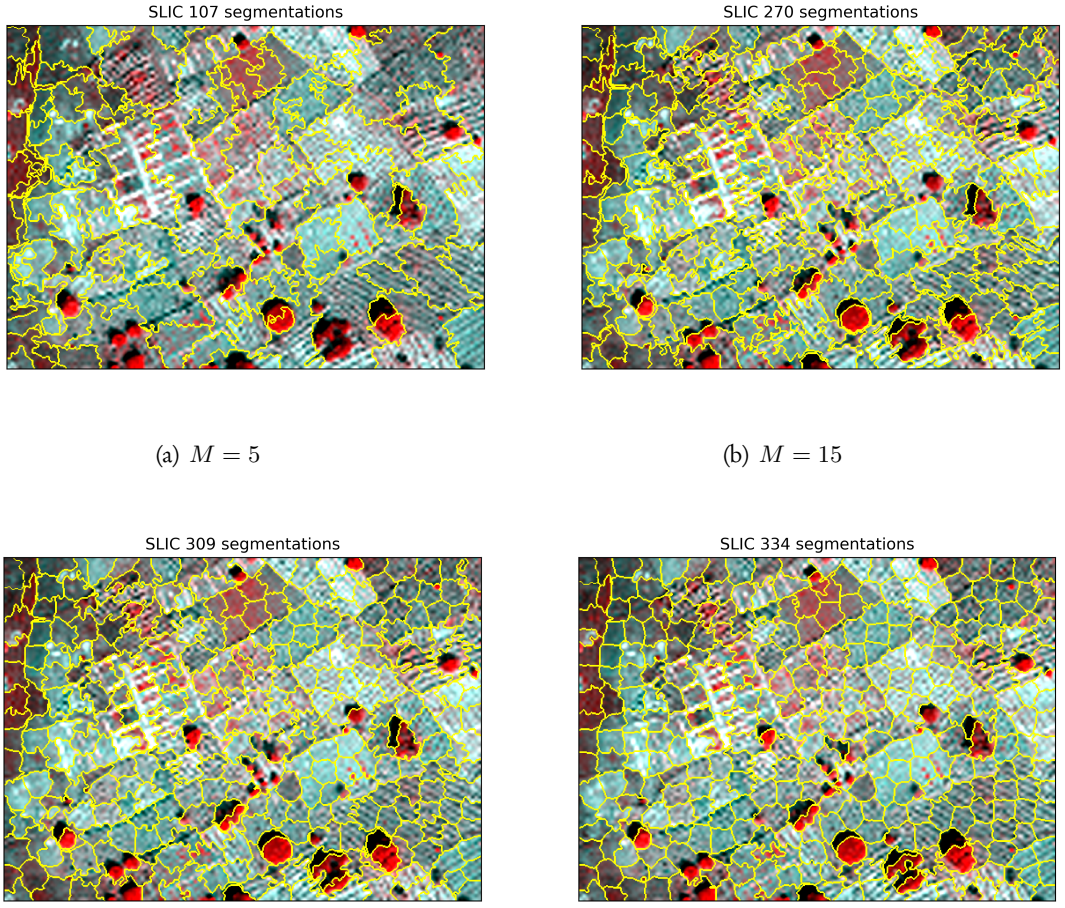
(a) $M = 5$ (b) $M = 15$

(c) $M = 25$ (d) $M = 35$

Figure 5.10: Boundaries for different values of M with fixed $K = 50$

Figure 5.10 and 5.11 show that as the value of M grows, the boundaries of segments become straighter. And because the condition to enforce connectivity is fixed, when M is small, many small segments are merged into the surrounding segments, which cause the number of superpixels from final results smaller than K. As the value of compactness M grows, the number of final superpixels grows obviously with it. At the same time, the weights of coordinate distance cost grows and the color distance cost does not change. Therefore the output superpixels are more and more decided by pixel position than by pixel color. This causes the final output's superpixels behave like regular grids.

As the authors of SLIC (Achanta et al., 2012) suggest, the range of compactness values is set to $[1, 40]$. In this project, the value of compactness is tested starting at 5 with increments of 5.

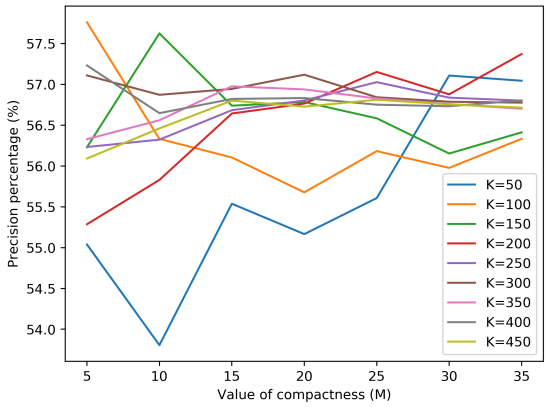


(a) $M = 5$ (b) $M = 15$

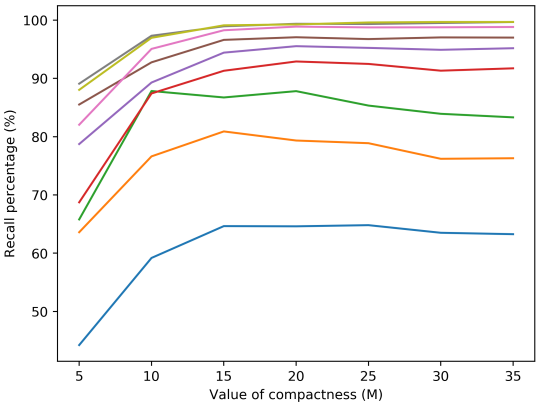
(c) $M = 25$ (d) $M = 35$

Figure 5.11: Boundaries for different values of M with fixed $K = 350$

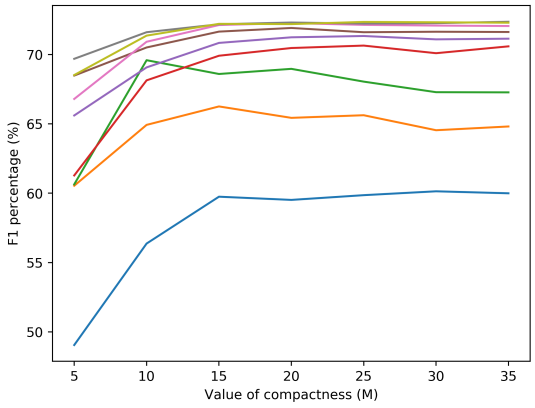
The curves of different M and fixed K of SLIC show an increase from 54 to 57. When $K = 50$, there is a small drop to 54 at $M = 10$ after that, while other results all stay around 56.5 with small fluctuations. These curves in figure 5.12(b) and (c) show that the value of compactness M does not influence the results when $K > 250$. For the ASA_s and ASA_g curves, the evaluation results have clear change from $M = 5$ to $M = 20$. For proper results for all methods, $M = 25$ is recognized as default value for more and further comparisons.



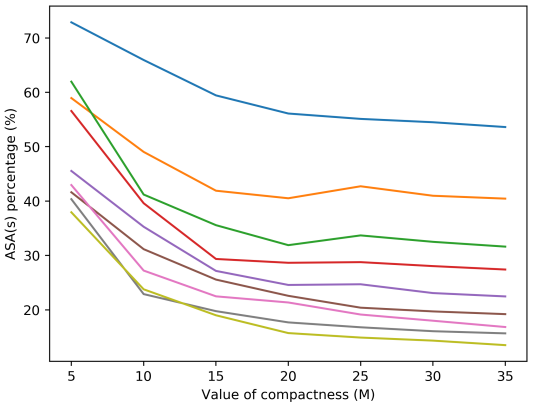
(a) Precision



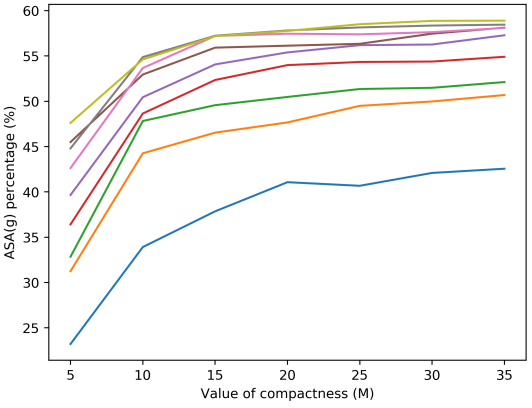
(b) Recall



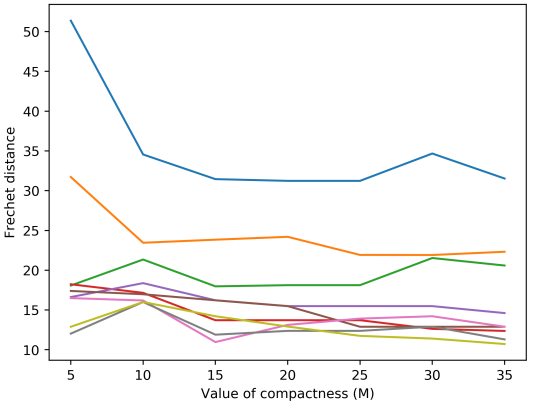
(c) $F1$



(d) ASA_s



(e) ASA_g

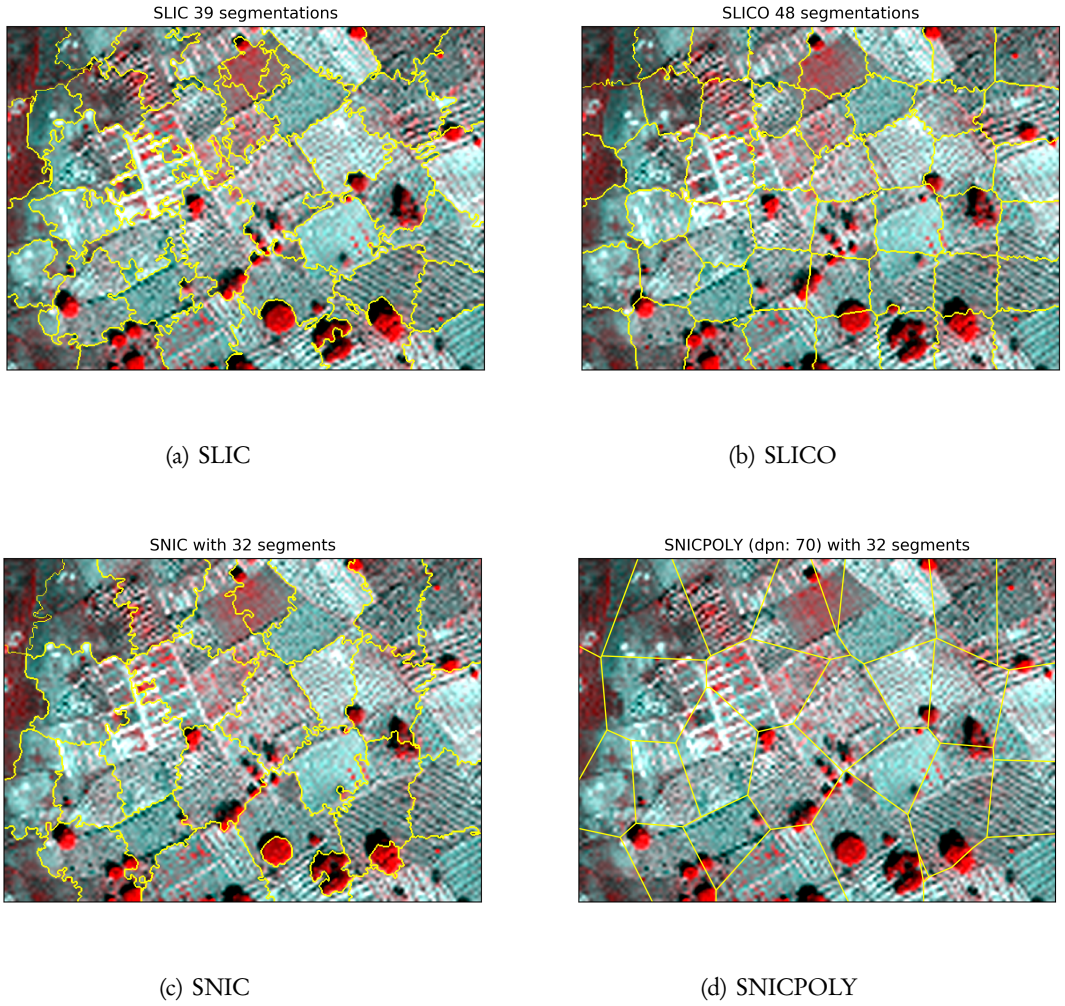


(f) Frechet distance

Figure 5.12: Evaluation results for different value of M with fixed K

5.5 COMPARISON BETWEEN THE SLIC, SLICO, SNIC, SNICPOLY ALGORITHMS

To compare the results of different runs of the SLIC algorithm, the values for M is fixed to 25, two values for K are tested at 50 and 150 respectively for better description of edges as discussed before. The results for small $K(50)$ are shown in figure 5.13 and the results for large $K(150)$ in figure 5.14.



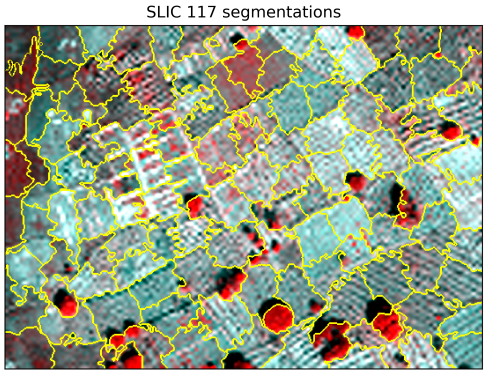
(a) SLIC (b) SLICO

(c) SNIC (d) SNICPOLY

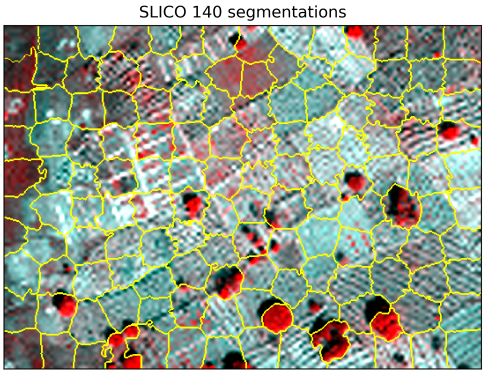
Figure 5.13: Comparison of SLIC, SLICO, SNIC and SNICPOLY with fixed number of superpixels $K = 50$, and compactness $M = 25$

As we knew at the start of this chapter, there are 34 ground truth parcels. When $K = 50$, four methods get the number of superpixels near the number of ground truth. SLIC determines some superpixels described as agricultural parcel. SLICO determines some edges connected to parcel boundaries, while SLICO results are more like regular grids. SNIC boundary description is better than that of SLICO. SLIC works better at edges with light and dark areas while SNIC works better at borders with textured and untextured areas than SLIC. SNICPOLY provides simplified boundaries of SNIC, which largely depend on the SNIC result.

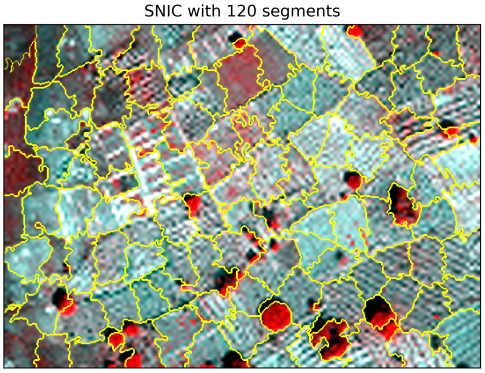
When K is much larger than the number of ground truth parcels (in this case, $K = 150$ which is three times larger), because of over-segmentation, the boundaries are basically extracted. SLIC



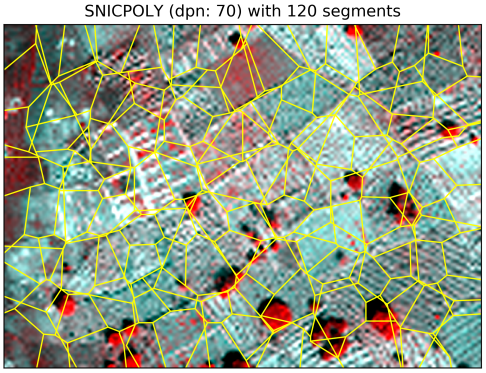
(a) SLIC



(b) SLICO



(c) SNIC



(d) SNICPOLY

Figure 5.14: Comparison of SLIC, SLICO, SNIC and SNICPOLY with fixed number of superpixels $K = 150$, and compactness $M = 25$

and SLICO do not extract edges where two sides have similar texture with a dark middle gap.

Comparing two conditions, results from SLIC are influenced by texture of agricultural field, especially when the parcel contains two types of crops and displays clear distinctions. The results from SLICO show more regular grids, have a balance between texture and the close connection with edges. The results from SNIC are more rugged than that of SLIC but are less affected by texture differences. Boundaries created by SNIC show a tendency of regular grids. And they also have good shapes of trees and tree shadow. The results from SNICPOLY are more artificial with fewer vertices and displays more straight lines than others.

Figures 5.15 show the evaluation with same M,

In terms of line evaluation, the precision, recall and $F1$ results for different version algorithms are presented in figures 5.15(a), (b) and (c) respectively.

In terms of area evaluation, ASA_s and ASA_g are used to measure different version algorithms and the results are presented in in figures 5.15(d) and (e) .

In terms of shape evaluation, Frechet distances are used to measure different version algorithms and the results are presented in in figures 5.15(f) .

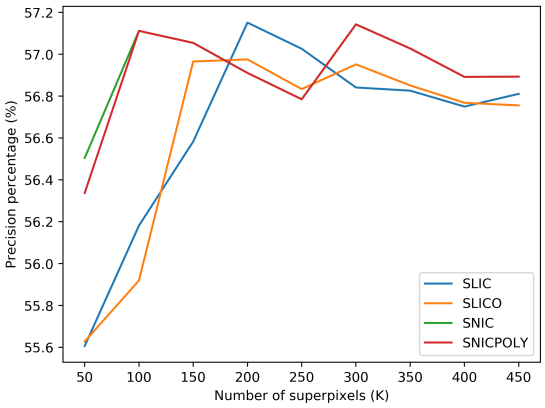
The precision percentage shows how many pixels are correctly recognized as boundary against the number of selected pixels. This percentage stays at the range of $[0, 1]$ and is expected as higher as possible. These curves show changes of precision with the increase of number of superpixels K and the increase of the value of compactness M.

For the curves for different K, the percentages of results from SLIC and SLICO methods increase from $K = 50$ to $K = 150$, while the percentages of SNIC and SNICPOLY peak at $K = 100$. Although there are differences between methods, generally all precision scores for different K and fixed M are around $(57 \pm 2)\%$. Based on these four curves, the result under $K = 150$ is considered as the best K.

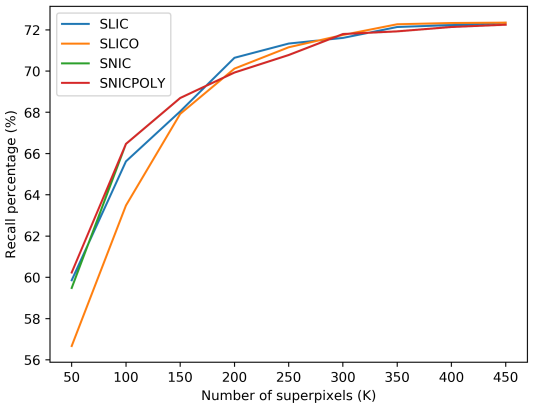
The recall percentage shows how many pixels are correctly selected as boundary under the number of ground truth boundary pixels. This percentage is in the range of $[0, 1]$. The higher the percentage, the better the result. As the number of superpixels K grows, the recall generally also grows. The percentages for different methods meet the same value and stay at it when $K \geq 350$. As for the change of value of compactness M, the results from SLIC give a rise from 82.5 to 98, while other results all stay around 98 with small fluctuations at the whole range of test M.

The $F1$ score is a harmonic average of precision and recall, which means this measurement combines the results of precision and recall. When $F1$ is higher, the result is considered as more ideal.

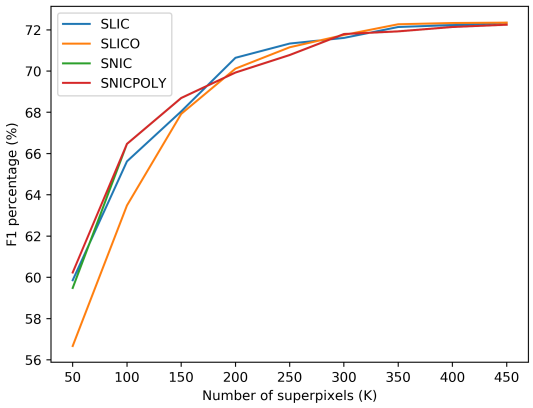
Because the precisions do not change much, the $F1$ scores are influenced much by the recall and the curves of $F1$ own the same tendency with the curves of recall. Based on the $F1$ scores, the best result is the same as that of recall under $K = 350$ and $M = 15$, which is shown in figure 5.14.



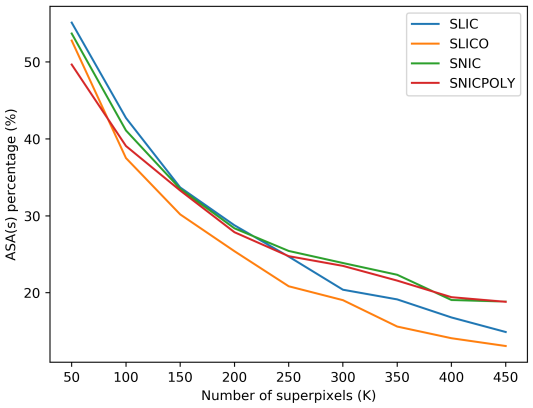
(a) Precision scores for different number of superpixels (K) with fixed value of compactness $M = 25$



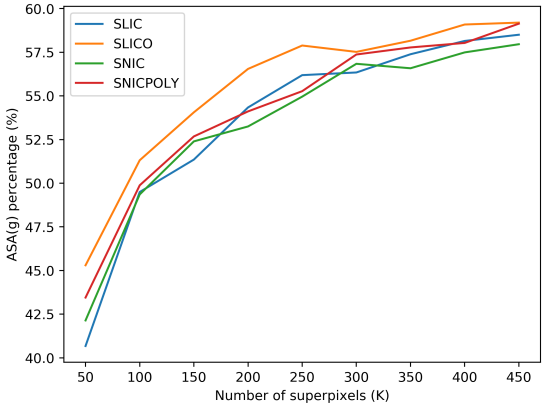
(b) Recall scores for different number of superpixels (K) with fixed value of compactness $M = 25$



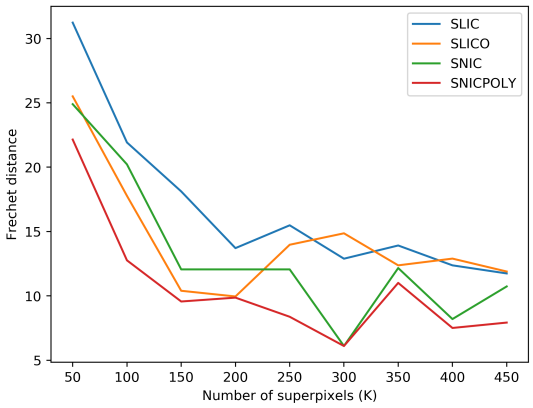
(c) $F1$ scores for different number of superpixels (K) with fixed value of compactness $M = 25$



(d) ASA_s for different number of superpixels (K) with fixed value of compactness $M = 25$

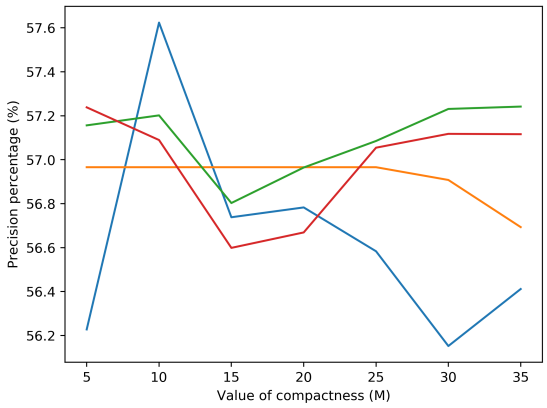


(e) ASA_g for different number of superpixels (K) with fixed value of compactness $M = 25$

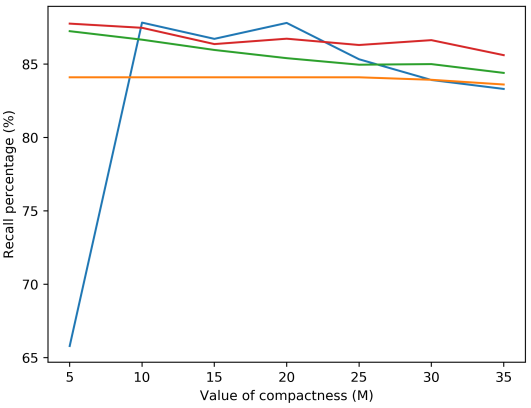


(f) Frechet distance for different number of superpixels (K) with fixed value of compactness $M = 25$

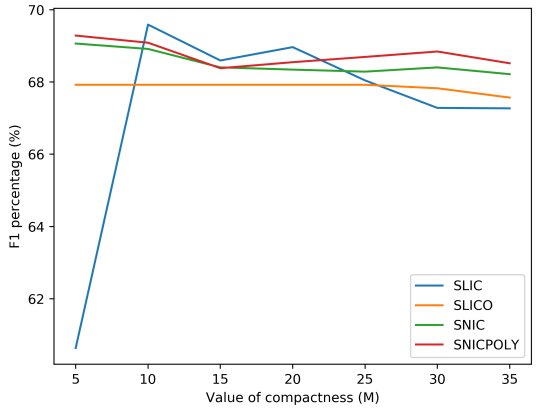
Figure 5.15: Evaluation graphs for different number of superpixels (K) with fixed value of compactness $M = 25$, images results shown in figure 5.13 and figure 5.14



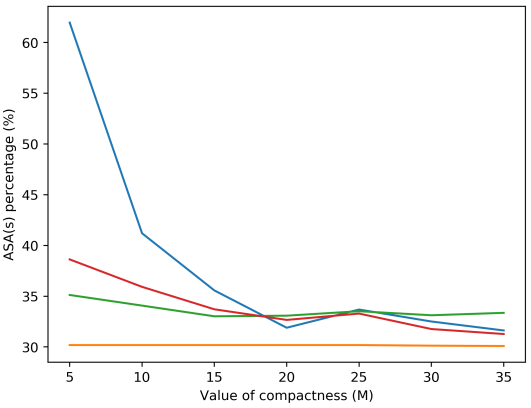
(a) Precision scores for different value of compactness (M) with fixed number of superpixels $K = 150$



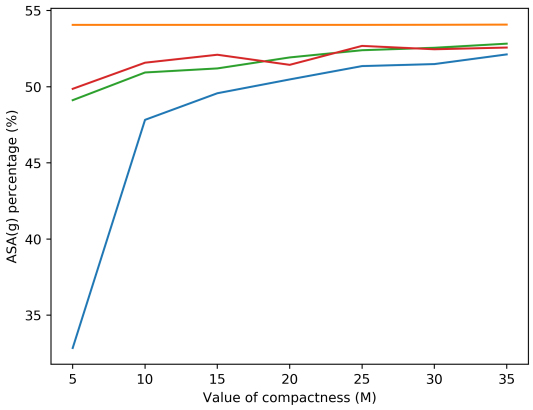
(b) Recall scores for different value of compactness (M) with fixed number of superpixels $K = 150$



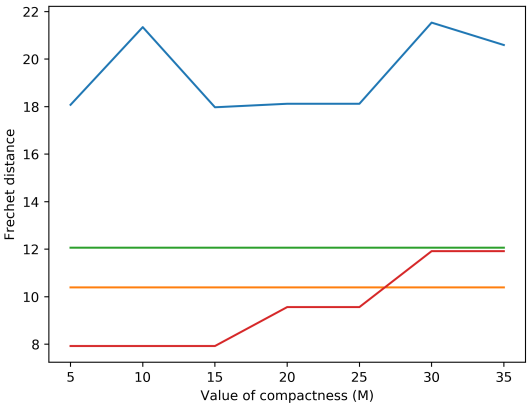
(c) $F1$ scores for different value of compactness (M) with fixed number of superpixels $K = 150$



(d) ASA_s for different value of compactness (M) with fixed number of superpixels $K = 150$



(e) ASA_g for different value of compactness (M) with fixed number of superpixels $K = 150$



(f) Frechet distance for different value of compactness (M) with fixed number of superpixels $K = 150$

Figure 5.16: Evaluation graphs for different value of compactness (M) with fixed number of superpixels $K = 350$, images results shown in figure 5.13 and figure 5.14

5.6 PARTIAL LEAST SQUARES METHOD EVALUATION

To use the PLS model, labels from ground truth data are used to build a training data set. For this project, the ground truth data is shown in figure 5.17. Labels per crop type range from 0 to 9 corresponding with the legend from top to bottom.

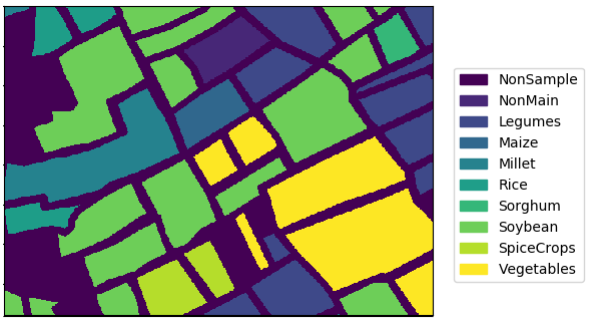


Figure 5.17: Ground truth labels

For each group, counts of ground truth pixels are presented in table 5.6.

Table 5.1: Basic information ground truth data

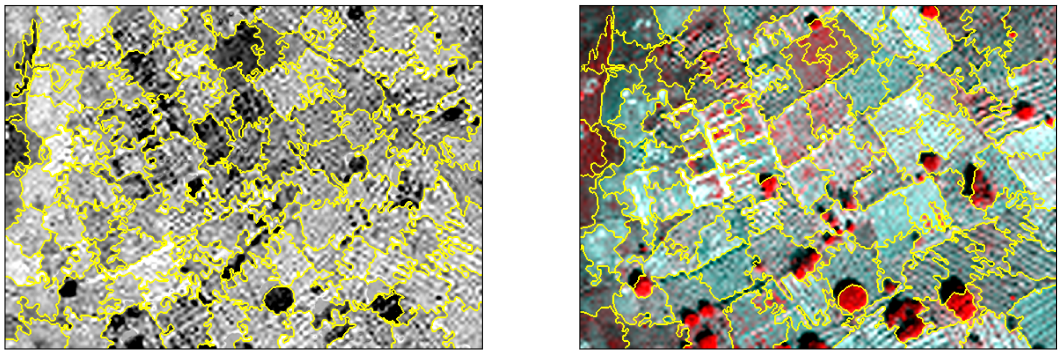
Label	Main type	Pixel count
1	Non-main	4304
2	Legumes	22333
3	Maize	4187
4	Millet	12714
5	Rice	5840
6	Sorghum	2064
7	Soybean	51896
8	Spice crops	6128
9	Vegetables	29802
	total	139268

To balance the training data, the same number of pixels are selected from each group to be part of the training set. And every pixel's eight bands data are used as training X and the pixel's corresponding label value is organized as its corresponding training Y to build a sample set to build a PLS model. This model is used to predict the value of each pixel in image I and obtain a predicted image I_P. We rescale I_P to a normal color range $[0, 255]$ for display as a gray image. We next use the predicted image I_P as red band, and copy it for green and blue bands to build a RGB image for use later in the SLIC algorithm.

The SLIC result of a PLS-predicted image I_P is shown in figure 5.18.

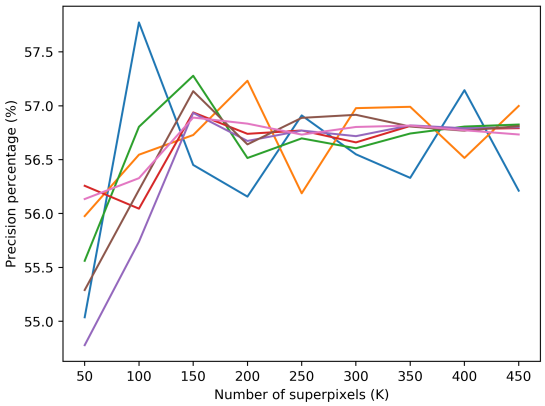
Although the PLS model gives more categorical values of what crop a pixel would be, the distinction between pixel values has diminished. This project assumed this method would decrease noise effects. In practice, it also decreases the differences between.

A range of K and M values are tested and evaluated, as shown in figure 5.19. Legends of all graphs are identical with the legend displayed in figure 5.19(c). From the ranges of different evaluation method (comparing figure 5.5, predicted image from PLS model get lower results than pan-sharpened image as discussed before.

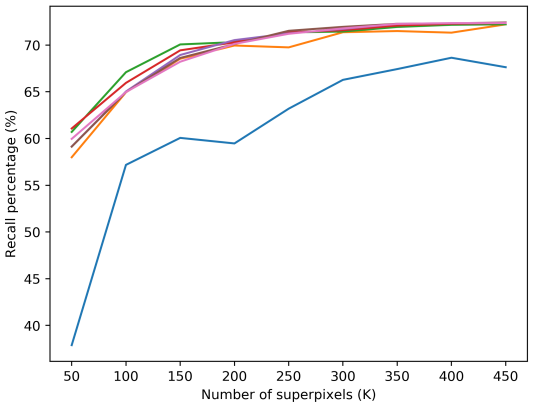


(a) PLS SLIC (b) Pan-sharpened SLIC

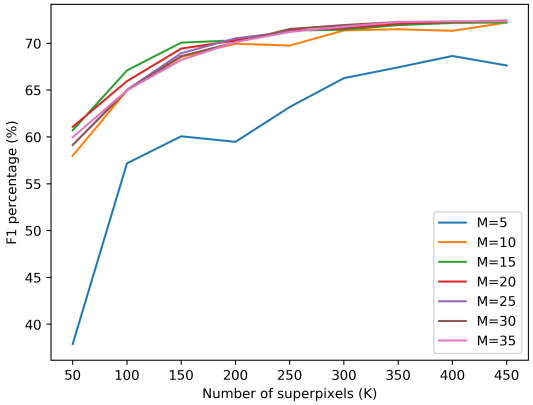
Figure 5.18: PLS SLIC result with fixed number of superpixels $K = 150$ and value of compactness $M = 15$



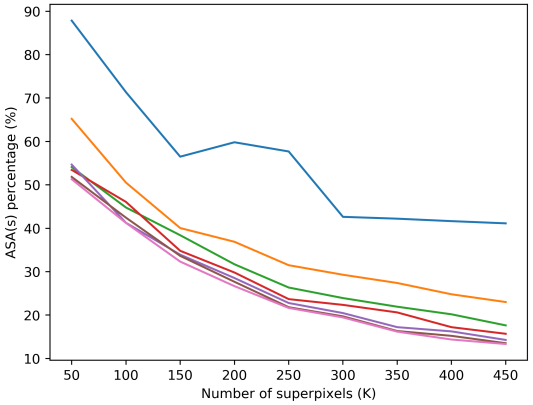
(a) Precision scores



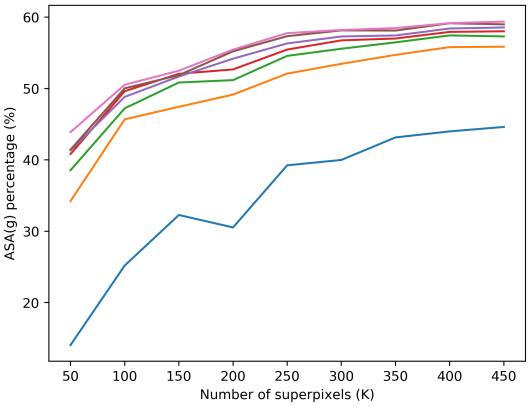
(b) Recall scores



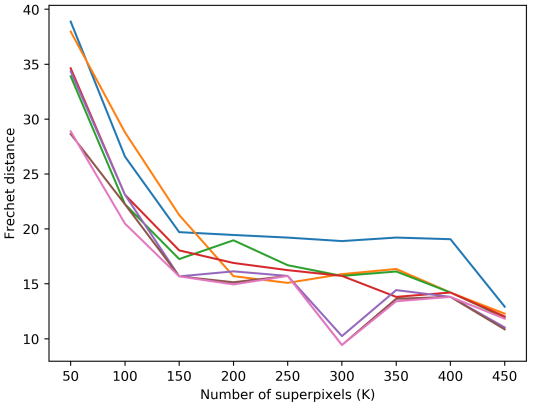
(c) $F1$ scores



(d) ASA_s



(e) ASA_g



(f) Frechet distance

Figure 5.19: PLS evaluation results for different value of K with fixed M

5.7 AGGREGATION

To see the comparison of different times of aggregation, two groups of parameter values are tested. The values of the number of superpixels K is kept fixed at 350 and 1350 respectively. For both, the value of compactness M is 15. The results are shown in figure 5.21. The segments after aggregation are displayed at the top of each figure.

According to the aggregation condition, the cost (based on equation 4.5.1) of each pair of connected segments gets accounted and is displayed in histogram of figure 5.20.

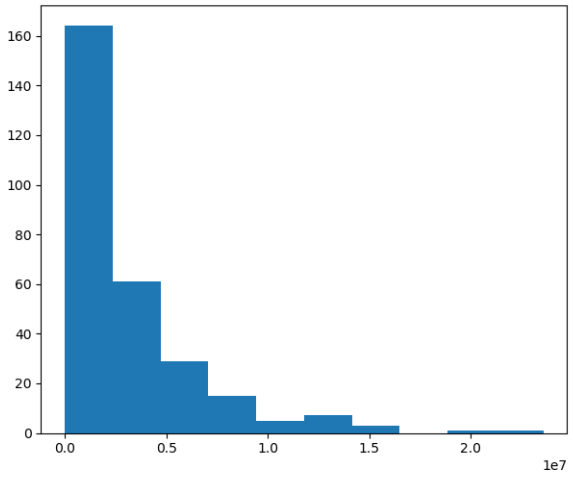
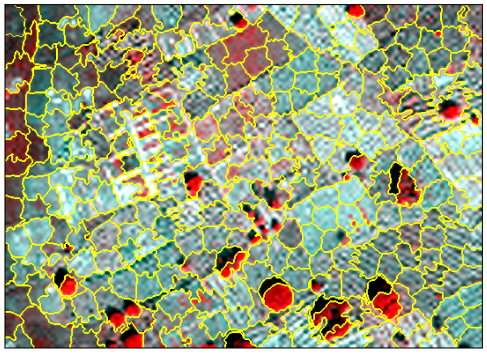


Figure 5.20: Histogram of aggregation cost

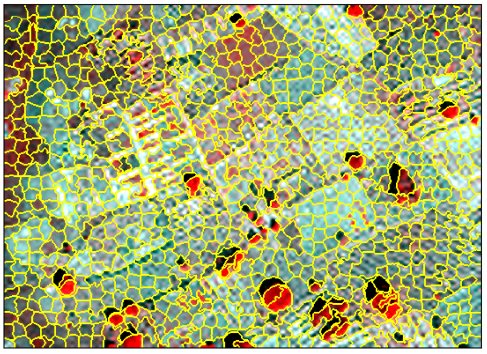
From this histogram, the threshold is tested from $2,000$ to $4,000$ with an increment $step = 500$. Threshold is tested and get the best result is obtained when $T = 3,000$. Aggregation results of multi-times are presented in figure 5.21.

From the first and second aggregation, trees and shadows of trees merge generally under small K condition ($K = 350$). When K is large, tree segments prefer to merge not with shadow segments, which is not wanted. More aggregation methods are needed to test for further work. The tree area influences the results of boundary delineation a lot. A more suitable way to treat tree area is also needed.

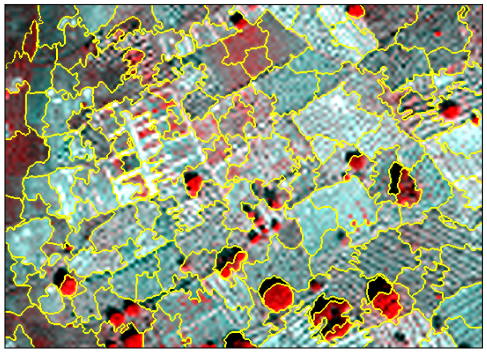
The percentages of precision, recall and $F1$ scores have a little grow at first aggregation and then decrease gradually. After aggregation, each pair of segments that meets the criteria (defined in chapter 4) is merged. Some line segments are dissolved. The ideal situation is that those dropped line segments are not a ground truth boundary. As the percentages go down, clearly some dropped line segments are actually true boundaries. At the same time, ASA_s and ASA_g both grow when more aggregation is allowed. This means that more and more area is labeled correctly within the ground truth parcels. Overall, a one time aggregation with a little larger K (compare with the number of ground truth parcels, in this case $K = 150$) obtains a better result.



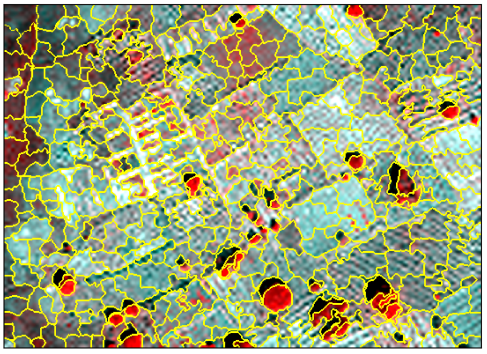
(a) $K = 350$, $M = 25$, 1 time aggregation, 89 segments



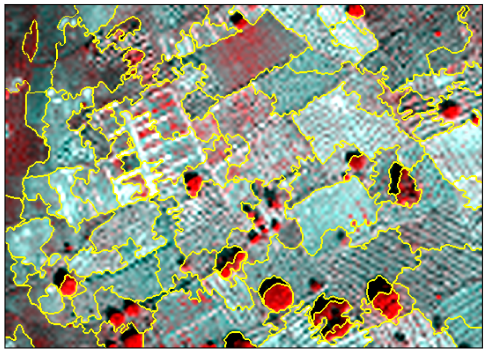
(b) $K = 1050$, $M = 25$, 1 time aggregation, 385 segments



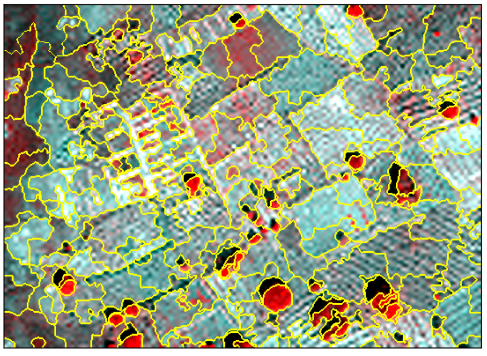
(c) $K = 350$, $M = 25$, 2 times aggregation, 30 segments



(d) $K = 1050$, $M = 25$, 2 times aggregation, 136 segments

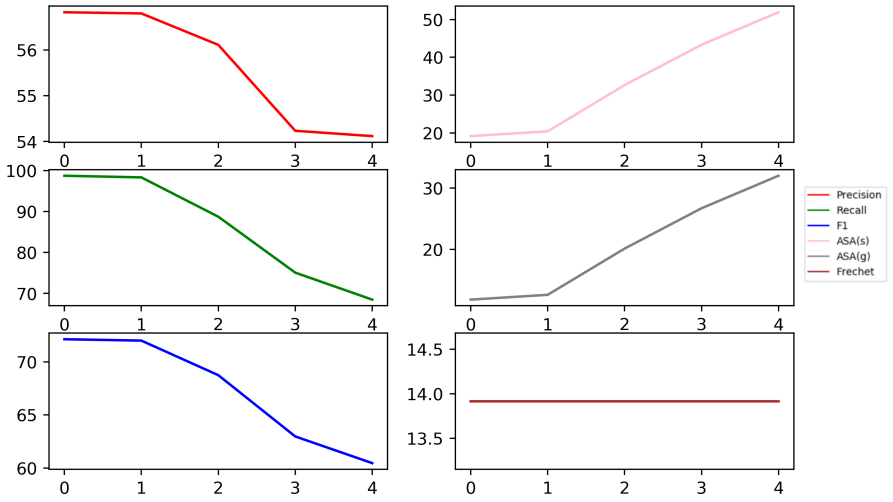


(e) $K = 350$, $M = 25$, 3 times aggregation, 13 segments

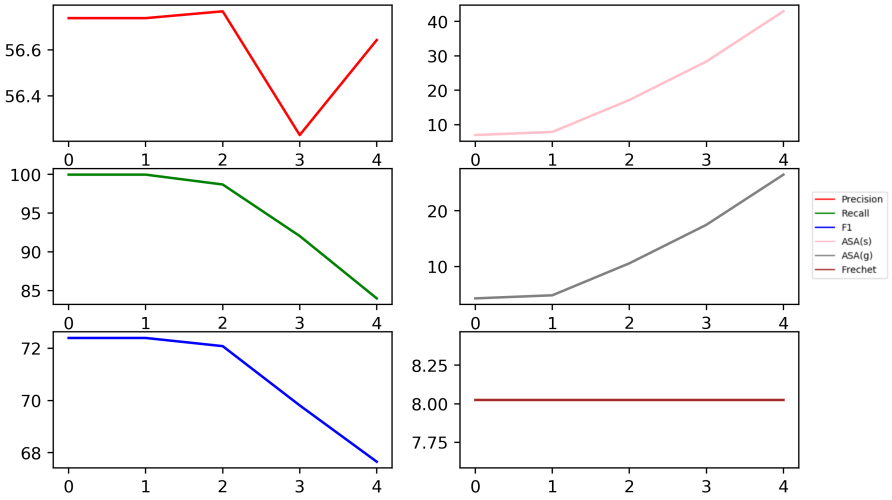


(f) $K = 1350$, $M = 25$, 3 times aggregation, 57 segments

Figure 5.21: Various times aggregation results comparison



(a) Aggregation start from $K = 350$, $M = 25$



(b) Aggregation start from $K = 1050$, $M = 25$

Figure 5.22: Evaluation of aggregation

Chapter 6
Discussion

This project aimed to discover the usefulness of superpixel generation for the delineation of agricultural parcels. At start, all we have is a multispectral image, a panchromatic image and ground truth vector data. Superpixel methods work on color images. Before superpixel generation, we organized to compare a multispectral false-color image, a panchromatic image, and a pan-sharpened image as input of our superpixel algorithms. We also proposed a PLS model to predict image value in hope for better delineation. In this project, we focused on the SLIC algorithm and tested three more algorithms, which are all related to SLIC. They are the SLICO, SNIC and SNICPOLY algorithms.

To measure the results of superpixel methods, three types of evaluation are used to compare results. These evaluation methods are the pixel-level precision, recall and $F1$ score, the area-level achievable segmentation accuracy and the shape-level Frechet distance between closed-ring parcel boundaries. These parameters are put together to pass a judgement on a results. During this project, superpixel algorithms were developed with which oversegmentation is common; oversegmentation gives better boundary recall scores. To achieve the eventual objective of proper parcel delineation, aggregation methods are considered that merge adjacent superpixels to decrease oversegmentation and retain proper boundaries at the same time. We discuss below three issues on our assumptions and results.

6.1 BASE DATA QUALITY

Because of time limitation, only one spectral and a panchromatic image are tested in this project. For further work, more images from different time should be researched. But in our case, this method works same in the different area from the sample area. This sample area is selected only because it has samples of every type of crop in ground truth data.

The ground truth data could be more accurate. There are large gaps between parcels and some of parcels are missed, which influence the results of evaluation.

6.2 CHOICES FOR K AND M PARAMETERS

These two parameters K and M are essential switch keys in this project. They are required by SLIC, SNIC and SNICPOLY. SLICO only requires the parameter K. SNICPOLY needs one more parameter dpn for the degree to simplify.

6.3 AGGREGATION METHOD

As the last section of chapter 5 shows, the aggregation results do not achieve the wanted target. The general condition that uses segment characterization with average scores shows some inappropriate pairing of segments to merge. We studied selected ground truth data to find out reasons.

From the labeled area, pixels are selected with values in eight bands to constitute a sample set. Next, we organized the sample set to have equal numbers of pixels per crop type. From the balanced sample set, we calculated the maximum, minimum, mean and standard deviation of each band per crop type. Lines were created for each statistical indicators to simulate average spectral curves and these are presented in figure 6.1.

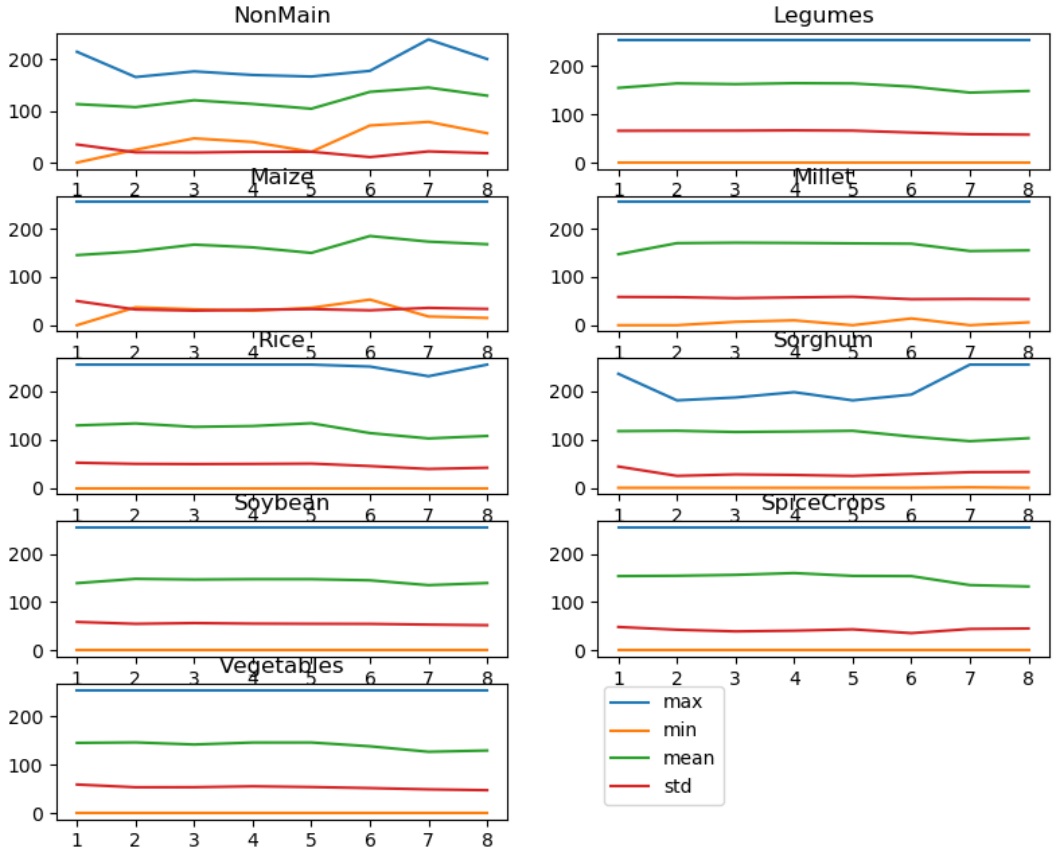


Figure 6.1: Statistical analysis per spectral band and per crop type in the sample set (x-axis: band no., y-axis: values)

As figure 6.1 shows, sorghum differs most notably from other crops. The maximum values of each sorghum band are respectively $236, 181, 187, 198, 181, 193, 255$ and 255, while others equal 255 except the maximum of band 7 of rice, which is 235. Despite that minimum values of all other crops equal 0, the minimum values of maize for different bands are $[0, 47, 33, 39, 46, 70, 40, 21]$ and the minimum values of band 3 and 4 of millet are 5 and 16. The whole image is stretched into range $[0, 255]$. Samples pixels are extracted from the whole image, it is grouped by its main crop types and expected to have individual features like spectral patterns. Unexpectedly, the samples from non-main type have the more diverse features.

Because the results after aggregation do not get expect accuracy, this means that distributions of sample data are not quite unified. The simple statistical data does not help much to find differences. We also developed univariate kernel density estimate curves of each band and each crop type as presented in figure 6.2 to display the range and density of values.

(Note: all the values mentioned are after rescaling to range $[0, 255]$.)

Figure 6.2 displays ranges of different bands and crop types to be similar. This makes using maximum and minimum values of the sample set useless. All these curves are simulated using a Gaussian Kernel. Although they are smooth, there are some interesting characteristics. For

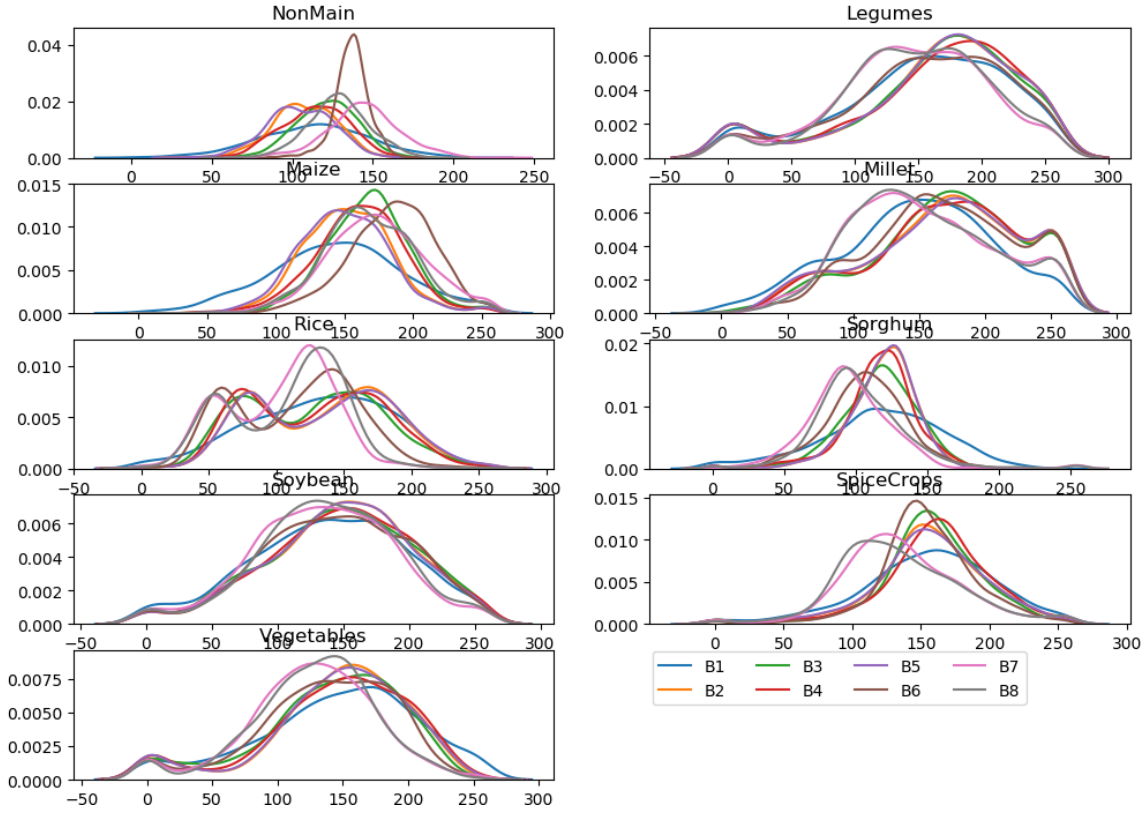


Figure 6.2: Univariate kernel density estimate per spectral band and per crop type in the sample set (X-axis: values, Y-axis: density)

legumes, the density of low value near 0 is a little larger than that when values near 50 in nearly all bands. Values near 100 to 200 of band 7 and 8 of legumes have high and similar density. The band 6 has similar curve with band 7 and 8 but is smaller and high pass section has a right shift. Curves of band 3, 4 and 5 have the highest density around 190. The curves of vegetables have identical features but also distinguishing ones. Another type of crop with an unusual pattern is rice. The curves of band 2, 3, 4 and 5 of rice have almost balanced two peaks, while the curves of 6, 7 and 8 have two unbalanced peaks. All right peaks have higher density than left ones. Other types of crop show different patterns also. Compare with average and standard deviation, density estimate curves could be more useful to distinguish each type of crop from each other. A proper method is wanted to organize density of each pair of segments to determine the cost function for aggregation. More work is needed to test this assumption and problems of computation memory and time complexity also need to be considered.

One probable reason of this problem is that in reality, the high-resolution image is clear enough and contains more information (bare soil, for instance) in its pixels. Even in one area of known crop at end of crop season, the crop does not cover every pixel. The pixel size is small enough to make some of them have pure features that are different from the known crop type. This explains that results obtained from a multispectral image are worse than those of pan-sharpened image results. This also causes the general spectral characterization to be useless. Smooth filters, like Gaussian filter could mitigate these pixel gaps while these filters also blur the image to some extent. In this context, when the plants are exuberant, the surface coverage is more complete, and the results are better. For the current situation, the density curves of each band of each type of crop land would have different patterns, but be helpful. Due to unchangeable spectrum of certain

objects, the density curves are like the weighted mean of two types of objects.

6.4 PARTIAL LEAST SQUARES MODEL

The partial least squares model has a good reputation but did not perform well in this project. More tests have been attempted to find out the reason.

First, the problem mentioned in the previous section also affects the PLS model. This is due to mixed ground truth pixels. The training data of the model contains noticeable noise.

Secondly, we use the PLS model as a pre-process. Each value is predicted pixel by pixel, which means each pixel area is dependent on surrounding pixels and the predicted values share the same range of Y available in the training set (introduced in section "Partial Least Squares Algorithm" of chapter 4). The predicted value range is dramatically smaller than the original range, which sharply decreases the distinction of each pixel, especially when using such predicted value as input of the clustering method SLIC. For SLIC, the color distance is an essential indicator. As separate pixel predicted values are calculated, this increases the noise in the color space, which leads to mistaken clustering and generates untrusted superpixels.

6.5 TREE AND SHADOW AREA

Figure 5.7 presents the boundaries with different number of superpixels K. As K grows, tree and shadow area have a greater impact on boundary delineation of agricultural area. It turns that superpixels deliver well-described $2D$ shapes of tree and shadows, as shown in figure 6.3.

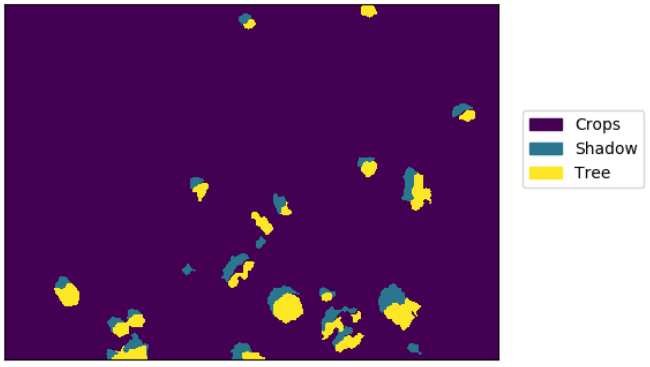


Figure 6.3: Tree and shadow area generated by SLIC

Because of time limitation, there is no further work on this, but two hypotheses have surfaced about how to handle tree and shadow area. The first hypothesis is to label tree and shadow area before boundary delineation. During superpixel generation, this area is put out of distance calculation. The second hypothesis is region filling (Antonio Criminisi, Patrick Perez, & Kentaro Toyama, 2004). The tree and shadow area is detected and refill with predicted values to pretend to be farmland.

Chapter 7
Conclusions and recommendations

7.1 CONCLUSIONS

This research project focuses on the performance of superpixel generation methods for the delineation of agricultural parcels using very high resolution satellite images. A workflow is designed to test and compare different methods of superpixel generation, with or without pre-processing. One multispectral and one panchromatic WorldView-3 image from November 20XX have been used in this project for boundary delineation.

The following answers to the research questions from chapter 1 were obtained:

a) Which image format works best for this project?

Which superpixel algorithm has the best performance delineation for SHA?

From visualization and evaluation results in chapter 5, we now know that pan-sharpened images work better than multispectral images, while the production technique of pan-sharpening does not matter much. The Brovery and Ihs methods present better texture lines. The sample mean method avoids pixel-level noise. All interpolation pan-sharpened methods give a better delineation of tree and shadow area. Nearest-neighbor and Lanczos methods deliver smoother images while linear and bicubic methods provide enhanced details.

This project compares four superpixel generation methods, SLIC, SLICO, SNIC, and SNICPOLY. When the parameter for the number of superpixels K is set near the real or expected number of parcels, SLIC works best. SLIC and SNIC methods delineate meaningful boundaries. SLICO delivers steady grids. SNICPOLY results largely depend on SNIC results. When K is large enough, although the image may become over-segmented, the boundaries of parcels are represented and this gives a chance for aggregation. In our case, with $K = 350$ (nearly ten times of the ground truth number), recall scores almost get 100 percent, and superpixels hand over enough results of boundary.

b) How does supervised feature reduction work with superpixels?

Does supervised feature reduction improve delineation quality of agricultural parcels?

An image is composed of PLS predicted values. Although this image decreases the effect of tree shadows, generally, the PLS-derived image performs not as good as pan-sharpened images. The percentage of selected boundary drops by 30%. Reasons are discussed in chapter 6.

c) How can one evaluate the results of agricultural boundary delineation in SHA?

This project delivers three methods to assess results. They are a pixel-level metric, an area size metric, and a shape metric. At pixel level, precision scores point out the percentage of selected boundary pixels that are actual ground truth boundary pixel. The recall scores depict the percentage of selected ground truth pixels over total ground truth pixels. $F1$ is a complex of precision score and recall score. At area level, ASA_s describes the proportion of correctly selected segments' area versus total ground truth area. ASA_g calculates the proportion of selected

ground truth area against the sum of each segment. For the shape metric, the Frechet distance is measured. This distance tells the maximum distance between segment boundary and ground truth boundary.

d) Under which conditions for K/M do Which methods work best?

Comprehensive assessment of all evaluation methods, for small K, when $K = 50$, $M = 25$, SNIC works best. For large K, when $K = 350$, $M = 25$, SLIC works best in our case.

e) How good is aggregation technique?

In our case, with $K = 1050$ (nearly thirty times of the ground truth number) and $M = 25$, recall scores are almost getting 100 percent and superpixels deliver enough boundary. When the threshold $T < 3000$, one-time aggregation provides the best result, which improve 2% accuracy in pixel level.

f) What is expected for this method to work in different areas?

This method works for areas with various types of crop. If the area is covered with lush plants, the area would be less heterogeneous and the results could be better.

7.2 RECOMMENDATIONS

The following points indicate the recommendations for future work:

* The area covered by trees needs to be separated from the agricultural area. The first way is to mask the tree area and no longer compare pixels in it during superpixel generation. This could need different repairs for each superpixel method, which could come with different consequences. Another considerable way is using region filling to remove trees (and shadow) and replace the value of such area with simulated values (Antonio Criminisi et al., 2004).

* Aggregation methods need more research to fix the condition to merge mixed segments. Probability density function could be consider to add into the condition equation to determine whether merge.

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Susstrunk, S. (2012). SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. on Pat. Anal. and Mach. Intel.*, *34*(1), 1–8. doi: 10.1109/tpami.2012.120

Achanta, R., & Süsstrunk, S. (2017, jul). Superpixels and polygons using simple non-iterative clustering. In *Proceedings - 30th ieee conference on computer vision and pattern recognition, cvpr 2017* (Vol. 2017-Janua, pp. 4895–4904). IEEE. Retrieved from http://ieeexplore.ieee.org/document/8100003/ doi: 10.1109/CVPR.2017.520

Antonio Criminisi, Patrick Perez, & Kentaro Toyama. (2004). Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, *13*(9), 1200–1212. Retrieved from www.csse.monash.edu.au/

Arbeláez, P., Pont-Tuset, J., Barron, J., Marques, F., & Malik, J. (2014). Multiscale combinatorial grouping. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *500*, 328–335. doi: 10.1109/CVPR.2014.49

Ballard, D. H. (1981, jan). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, *13*(2), 111–122. Retrieved from https://www.sciencedirect.com/science/article/pii/0031320381900091 doi: 10.1016/0031-3203(81)90009-1

Bayer. (2018). *Smallholder Farming: Small Land, Large Impact.* Retrieved 2019-01-02, from https://www.cropscience.bayer.com/en/crop-science/smallholder-farming

Benediktsson, J. A., & Ghamisi, P. (2015). *Spectral-Spatial Classification of Hyperspectral Remote Sensing Images.* Retrieved from https://books.google.nl/books?hl=en{\&}lr={\&}id=TtnRCgAAQBAJ{\&}oi=fnd{\&}pg=PR7{\&}dq=Spectral-Spatial+Classification+of+Hyperspectral+Remote+Sensing+Images{\&}ots=fV95G-Bzxy{\&}sig=aSIrBRNA{_}RVVSJwhEOdOJDW5gb0{\&}redir{_}esc=y{\#}v=onepage{\&}q=Spectral-Spatial{\%}2520Classification{\%}2520

Blaschke, T. (2010, jan). *Object based image analysis for remote sensing* (Vol. 65) (No. 1). Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/S0924271609000884 doi: 10.1016/j.isprsjprs.2009.06.004

Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, *40*(1), 16–28. Retrieved from http://dx.doi.org/10.1016/j.compeleceng.2013.11.024 doi: 10.1016/j.compeleceng.2013.11.024

Chen, L., Ma, M., Jing, N., Guo, N., & Xiong, W. (2017, oct). An Efficient Query Algorithm for Trajectory Similarity Based on Fréchet Distance Threshold. *ISPRS International Journal of Geo-Information*, *6*(11), 326. Retrieved from http://www.mdpi.com/2220-9964/6/11/326 doi: 10.3390/ijgi6110326

Collins, R. (n.d.). Mean-shift blob tracking through scale space. In *2003 ieee computer society conference on computer vision and pattern recognition, 2003. proceedings.* (Vol. 2, pp. II–234–40). IEEE Comput. Soc. Retrieved from http://ieeexplore.ieee.org/document/1211475/ doi: 10.1109/CVPR.2003.1211475

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on pattern analysis and machine intelligence*, *24*(5), 603–619.

Dossou-Yovo, E. R., Kouyaté, A. M., Sawadogo, T., Ouédraogo, I., Bakare, O. S., & Zwart, S. J.

(2018, aug). A geospatial database of drought occurrence in inland valleys in Mali, Burkina Faso and Nigeria. *Data in Brief*, *19*, 2008–2014. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S235234091830756X doi: 10.1016/j.dib.2018.06.105

Douglas, D. H., & Peucker, T. K. (1973, dec). Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. *The International Journal for Geographic Information and Geovisualization*, *10*(2), 15–28. Retrieved from https://utpjournals.press/doi/10.3138/FM57-6770-U75U-7727 doi: 10.1002/9780470669488.ch2

Fan, S., Brzeska, J., Keyzer, M., & Halsema, A. (2013). From Subsistence to Profit: Transforming Smallholder Farms. *IFPRI*. Retrieved from http://dx.doi.org/10.2499/9780896295582 doi: 10.2499/9780896295582

FAO. (2010). *Part two agricultural census items-concepts, definitions and tabulations, Chapter 11 concepts and definitions of agricultural census items* (Tech. Rep.). Retrieved from http://www.fao.org/fileadmin/templates/ess/documents/world{_}census{_}of{_}agriculture/chapter11{_}r7.pdf

Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, *59*(2), 167–181. Retrieved from http://people.cs.uchicago.edu/{~}pff/papers/seg-ijcv.pdf doi: 10.1023/B:VISI.0000022288.19776.77

Food and Agriculture Organization of the United Nations. (2005). *A system of integrated agricultural censuses and surveys*. Food and Agriculture Organization of the United Nations. Retrieved from http://www.fao.org/docrep/009/a0135e/A0135E00.htm{\#}TOC

García-Pedrero, A., Gonzalo-Martín, C., & Lillo-Saavedra, M. (2017, apr). A machine learning approach for agricultural parcel delineation through agglomerative segmentation. *International Journal of Remote Sensing*, *38*(7), 1809–1819. Retrieved from https://www.tandfonline.com/doi/full/10.1080/01431161.2016.1278312 doi: 10.1080/01431161.2016.1278312

Geladi, P., & Kowalski, B. R. (1986). Partial least-squares regression: a tutorial. *Analytica chimica acta*, *185*, 1–17.

Gollin, D. (2014). Smallholder agriculture in Africa: An overview and implications for policy. , 14640. Retrieved from http://pubs.iied.org/pdfs/14640IIED.pdf

Grompone Von Gioi, R., Jakubowicz, J., Morel, J. M., & Randall, G. (2010, apr). LSD: A fast line segment detector with a false detection control. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(4), 722–732. Retrieved from http://ieeexplore.ieee.org/document/4731268/ doi: 10.1109/TPAMI.2008.300

Haenlein, M., & Kaplan, A. M. (2004, nov). A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics*, *3*(4), 283–297. Retrieved from http://www.tandfonline.com/doi/abs/10.1207/s15328031us0304{_}4 doi: 10.1207/s15328031us0304_4

Haris, K., Efstratiadis, S. N., Maglaveras, N., & Katsaggelos, A. K. (1998). Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, *7*(12), 1684–1699. Retrieved from https://www.researchgate.net/publication/5576266 doi: 10.1109/83.730380

Huang, S. H. (2015). Supervised feature selection: A tutorial. *Artificial Intelligence Research*, *4*(2). Retrieved from http://www.sciedupress.com/journal/index.php/air/article/view/6218 doi: 10.5430/air.v4n2p22

Inan, H. I., Sagris, V., Devos, W., Milenov, P., van Oosterom, P., & Zevenbergen, J. (2010, dec). Data model for the collaboration between land administration systems and agricultural land parcel identification systems. *Journal of Environmental Management*,

91(12), 2440–2454. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0301479710002112 doi: 10.1016/j.jenvman.2010.06.030

ITC. (2014). *STARS Project*. Retrieved 2019-01-02, from https://www.stars-project.org/en/

Khalid, S., Khalil, T., & Nasreen, S. (2014). A survey of feature selection and feature extraction techniques in machine learning. *2014 Science and Information Conference*, 372–378. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6918213 doi: 10.1109/SAI.2014.6918213

Levinshtein, A., Stere, A., Kutulakos, K. N., Fleet, D. J., Dickinson, S. J., & Siddiqi, K. (2009, dec). TurboPixels: Fast superpixels using geometric flows. In *Ieee transactions on pattern analysis and machine intelligence* (Vol. 31, pp. 2290–2297). Retrieved from http://ieeexplore.ieee.org/document/4912213/ doi: 10.1109/TPAMI.2009.96

Li, H., Xiong, P., An, J., & Wang, L. (2018). Pyramid Attention Network for Semantic Segmentation. , 1–13. Retrieved from http://arxiv.org/abs/1805.10180 doi: 10.1111/j.1532-5415.2011.03560.x

Lowder, S. K., Skoet, J., & Raney, T. (2016, nov). The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development*, *87*, 16–29. Retrieved from https://www.sciencedirect.com/science/article/pii/S0305750X15002703 doi: 10.1016/j.worlddev.2015.10.041

Lu, S., Oki, K., Shimizu, Y., & Omasa, K. (2007). Comparison between several feature extraction/classification methods for mapping complicated agricultural land use patches using airborne hyperspectral data. *International Journal of Remote Sensing*, *28*(5), 963–984. doi: 10.1080/01431160600771561

Marr, D., & Hildreth, E. (1980). Theory of Edge Detection. *Proc. R. Soc. Lond. B*, *207*, 187–217. Retrieved from https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.1980.0020 doi: 10.1098/rspb.1980.0020

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Proceedings of the IEEE International Conference on Computer Vision*, *2*, 416–423. Retrieved from http://ieeexplore.ieee.org/document/937655/ doi: 10.1109/ICCV.2001.937655

Mokrzycki, W., & M, S. (2012). New version of Canny edge detection algorithm. (May), 533–540.

Moore, A. P., Prince, S. J., Warrell, J., Mohammed, U., & Jones, G. (2008, jun). Superpixel lattices. In *26th ieee conference on computer vision and pattern recognition, cvpr* (pp. 1–8). IEEE. Retrieved from http://ieeexplore.ieee.org/document/4587471/ doi: 10.1109/CVPR.2008.4587471

Nagaraja, V. K., & Abd-Almageed, W. (2015, jul). Feature Selection using Partial Least Squares Regression and Optimal Experiment Design. In *Proceedings of the international joint conference on neural networks* (Vol. 2015-Septe, pp. 1–8). IEEE. Retrieved from http://ieeexplore.ieee.org/document/7280341/ doi: 10.1109/IJCNN.2015.7280341

Nikhil, R., & Sankar, K. (1993, sep). A review on image segmentation techniques. *Pattern Recognition*, *26*(9), 1277–1294. Retrieved from https://www.sciencedirect.com/science/article/pii/003132039390135J doi: 10.1016/0031-3203(93)90135-J

Ogungbile, A., Tabo, R., Duivenbooden, N. V., & Debrah, S. (2012, may). Analysis of constraints to agricultural production in the Sudan Savanna zone of Nigeria using multi-scale characterization. *NJAS wageningen journal of life sciences*, *59*(1), 53–60. Retrieved from https://library.wur.nl/ojs/index.php/njas/article/view/494 doi: 10.1016/j.njas.2012.02.001

Opencv dev team. (2019). *Geometric Image Transformations — OpenCV 2.4.13.7 documentation.* Retrieved 2019-02-02, from https://docs.opencv.org/2.4/modules/imgproc/doc/geometric{_}transformations.html

Padwick, C., Deskevich, M., Pacifici, F., & Smallwood, S. (2010). WorldView-2 pan-sharpening. *ASPRS 2010*. Retrieved from http://www.asprs.org/wp-content/uploads/2013/08/Padwick.pdf

Reso, M., Jachalsky, J., Rosenhahn, B., & Ostermann, J. (2013). Temporally consistent superpixels. In (pp. 385–392). Retrieved from https://www.cv-foundation.org/openaccess/content{_}iccv{_}2013/html/Reso{_}Temporally{_}Consistent{_}Superpixels{_}2013{_}ICCV{_}paper.html doi: 10.1109/ICCV.2013.55

Rydberg, A., & Borgefors, G. (2001). Integrated method for boundary delineation of agricultural fields in multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(11), 2514–2520. Retrieved from http://ieeexplore.ieee.org/document/964989/ doi: 10.1109/36.964989

Safdar, M., Cui, G., Kim, Y. J., & Luo, M. R. (2017). Perceptually uniform color space for image signals including high dynamic range and wide gamut. *Optics Express*, *25*(13), 15131. Retrieved from https://doi.org/10.1364/OE.25.015131https://www.osapublishing.org/abstract.cfm?URI=oe-25-13-15131 doi: 10.1364/OE.25.015131

Sheikhpour, R., Sarram, M. A., Gharaghani, S., & Chahooki, M. A. Z. (2017, apr). A Survey on semi-supervised feature selection methods. *Pattern Recognition*, *64*, 141–158. Retrieved from https://www.sciencedirect.com/science/article/pii/S0031320316303545 doi: 10.1016/J.PATCOG.2016.11.003

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(8), 888–905. Retrieved from http://ieeexplore.ieee.org/document/868688/ doi: 10.1109/34.868688

Star, S. L. (2010, sep). This is not a boundary object: Reflections on the origin of a concept. *Science Technology and Human Values*, *35*(5), 601–617. Retrieved from http://journals.sagepub.com/doi/10.1177/0162243910377624 doi: 10.1177/0162243910377624

Stutz, D., Hermans, A., & Leibe, B. (2018, dec). Superpixels: An evaluation of the state-of-the-art. *Computer Vision and Image Understanding*, *166*, 1–27. Retrieved from http://arxiv.org/abs/1612.01601http://dx.doi.org/10.1016/j.cviu.2017.03.007 doi: 10.1016/j.cviu.2017.03.007

United Nations. (2015). *Food security and nutrition and sustainable agriculture: Sustainable Development Knowledge Platform.* Retrieved 2018-08-27, from https://sustainabledevelopment.un.org/topics/foodagriculture

Vincent, L., Vincent, L., & Soille, P. (1991, jun). Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(6), 583–598. Retrieved from http://ieeexplore.ieee.org/document/87344/ doi: 10.1109/34.87344

Wang, H., & Sadiq, S. (2013). An Effectiveness Study on Trajectory Similarity Measures(ADC 2013).pdf. In (Vol. 137, pp. 13–22). Retrieved from http://delivery.acm.org/10.1145/2530000/2525418/p13-wang.pdf?ip=130.89.137.143{\&}id=2525418{\&}acc=PUBLIC{\&}key=0C390721DC3021FF.7DEDEACE9AC2380A.4D4702B0C3E38B35.4D4702B0C3E38B35{\&}{_}{_}acm{_}{_}=1550654124{_}3b26903256832c58e23bd392b7dfa30fhttp://dl.acm.org/citation.cfm

Wold, S., Sjöström, M., & Eriksson, L. (2001, oct). PLS-regression: A basic tool of chemometrics. In *Chemometrics and intelligent laboratory systems* (Vol. 58, pp. 109–130). El-

sevier. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169743901001551 doi: 10.1016/S0169-7439(01)00155-1

X. Ren, & J. Malik. (2003). Learning a classification model for segmentation. *Proceedings Ninth IEEE International Conference on Computer Vision*, *1*(c), 10–17 vol.1. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1238308 doi: 10.1109/ICCV.2003.1238308

Zhang, G., Jia, X., & Hu, J. (2015, nov). Superpixel-based graphical model for remote sensing image mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(11), 5861–5871. Retrieved from http://ieeexplore.ieee.org/document/7119598/ doi: 10.1109/TGRS.2015.2423688

Zhao, Z., & Liu, H. (2007, apr). Semi-supervised feature selection via spectral analysis. *Proceedings of the 7th SIAM International …*, 641–646. Retrieved from https://epubs.siam.org/doi/10.1137/1.9781611972771.75http://www.public.asu.edu/{~}huanliu/papers/ssfs.pdf doi: 10.1137/1.9781611972771.75