

**VISIBLE CADASTRAL BOUNDARY  
EXTRACTION USING VHR REMOTE  
SENSING IMAGES:  
A DEEP LEARNING APPROACH**

BEDRU WUDYE TAREKE

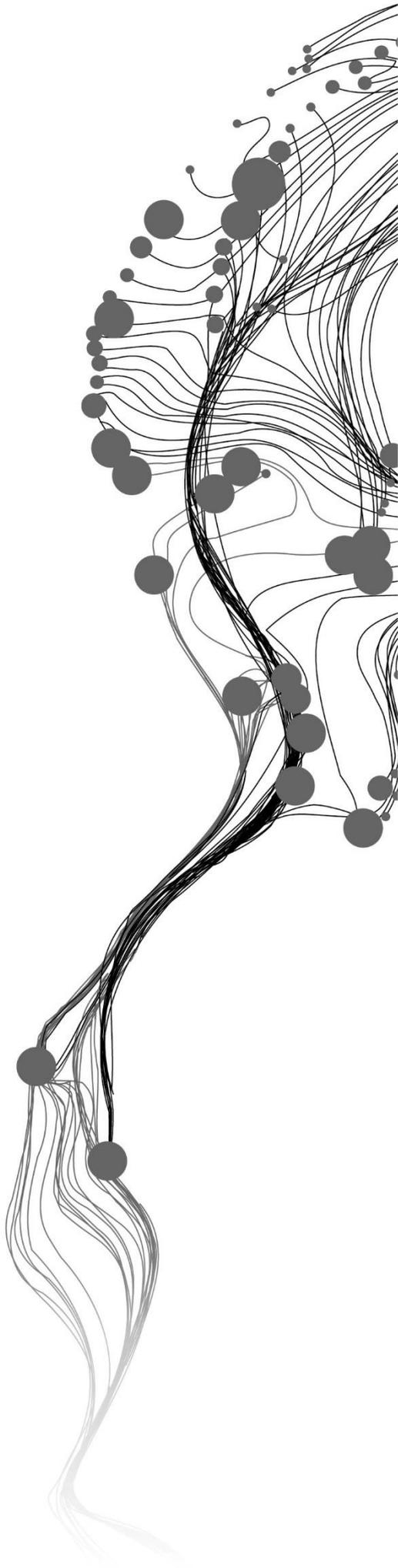
July 2022

SUPERVISORS:

Dr. M.N, Koeva

Prof. dr. C. Persello





# **VISIBLE CADASTRAL BOUNDARY EXTRACTION USING VHR REMOTE SENSING IMAGES: A DEEP LEARNING APPROACH**

**BEDRU WUDYE TAREKE**

Enschede, The Netherlands, July 2022

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: Geo-Information Management for Land Administration

**SUPERVISORS:**

Dr. M.N Koeva

Prof. dr. C. Persello

**THESIS ASSESSMENT BOARD:**

Prof. dr. Ir. J.A Zevenbergen (Chair)

Dr. Simon Hull (External Examiner, University of Cape Town)

#### 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

The United Nations sustainable development goal (SDG), particularly indicator 1.4.2, promotes tenure security for all, special attention given to the poor and vulnerable groups. Despite this, majority of the global population lacks access to formal land registration systems. It is exacerbated by the expensive and time consuming nature of conventional systems of land registration such use of high precision ground survey methods. Due to its flexibility and affordability, several countries have recently adopted fit-for-purpose land administration (FFP-LA) approach. FFP-LA promotes use of country wide aerial orthophoto, incremental improvement of measurement accuracy and using general boundary principle to speed up the process of land rights mapping. In recent years, the feasibility of using image-based mapping and automatic feature extraction techniques for (semi) automated cadastral boundary extraction has been actively investigated as an alternative to the traditional mapping methods.

Many GIS applications, including cadastral mapping, require vector representation of spatial objects. Although deep learning networks such as FCN have achieved remarkable result in segmenting remote sensing images, the final map of the segmentation output is in raster format. Inspired by FFP-LA, this research has investigated the usability of deep learning methods to extract cadastral boundaries from very high resolution (VHR) remote sensing images directly in a vector polygon format. The proposed method utilizes FCN based network specifically UNET as a backbone to perform the segmentation task on the input image followed by the frame field method which simultaneously learns directional information of each pixel on the image. Through multi-task learning, frame field output improves the quality of deep segmentation model while providing structural information to facilitate the polygonization process. The polygonization method inspired by active contour model (ACM) takes the segmentation mask and frame field information as input and iteratively optimizes edges and corners to align seamlessly with the reference data.

The experiments were conducted in the rural part of Ethiopia where agriculture is the predominant land use. The proposed method was evaluated on two different orthophotos based on images taken from UAV and aircraft platforms with 0.11 and 0.29 cm. resolution, respectively. The results are evaluated and reported using PoLiS and IoU metric. Polygons predicted on UAV orthophotos have higher similarity with the reference polygon than polygons predicted on aerial orthophotos with PoLiS distance of 2.81 and 8.64, respectively. Furthermore, a higher mean IoU of 0.84 was achieved on polygons predicted using UAV orthophotos compared to predictions on aerial orthophotos which is 0.79. Frame field based model predictions delivered simplified polygons with regular edge and corner compared to standard segmentation model. In addition, the model is tested in a different geographical area using aerial orthophoto to test its transferability. mIoU of 0.67 and PoLiS distance of 7.11 is achieved which is slightly lower than the quality of polygons predicted using UAV orthophoto. The accuracy of the model was influenced by many factors such as high crop variability inside one field, existence of invisible boundaries, confusing features such as water ways and terraces that crosses the parcel, date of image acquisition and the quality of reference cadastral data.

In conclusion, the proposed method shows the possibility of utilizing deep learning methods for extracting cadastral boundaries from VHR remote sensing images in a vector polygon format that can be directly used in mapping applications with little post-processing.

Keywords: frame field learning, deep learning, FCN, UNET, cadastral mapping, UAV, orthophoto, FFP-LA

## ACKNOWLEDGEMENTS

My first and foremost gratitude goes out to Lord the Almighty, for granting me countless blessings, knowledge, and health that allowed me to accomplish this thesis.

I would like to thank the following people, without whom I would not have been able to complete this research, and without whom I would not have made it through my master's degree. Special thanks to my supervisors Dr. M.N Koeva and Prof.dr. C. Persello, who read my numerous revisions and helped make some sense of the confusion. It was extremely fortunate for me to have a supervisor who cared so much about me and my work, and who always responded promptly to my queries and questions.

I extend my gratitude to Mr. Bezuaem Bekele who always wanted to see me succeed and gave me numerous opportunities since the beginning of my career. Thank you for the tremendous support you provided. I am indebted to you.

Thanks to the University of Twente for awarding me a fellowship, providing me with the financial means to complete my studies. And finally, my biggest thanks to my family, especially my big brother who endured this long journey with me, always offering support and love.

# TABLE OF CONTENTS

---

1.	INTRODUCTION.....	7
1.1.	Background.....	7
1.2.	Research problem.....	10
1.3.	Research objective and questions.....	11
1.4.	Conceptual framework.....	12
1.5.	Ethical considerations.....	12
1.6.	Structure of the thesis.....	13
1.7.	Summary.....	13
2.	LITERATURE REVIEW.....	14
2.1.	Cadastral boundaries & mapping methods.....	14
2.2.	Image-based automatic boundary extraction.....	15
2.3.	Fully Convolutional Neural Networks.....	15
2.4.	UNET.....	16
2.5.	SegNet.....	17
2.6.	Summary.....	17
3.	RESEARCH METHODOLOGY.....	18
3.1.	General workflow.....	18
3.2.	Study area & datasets.....	19
3.3.	Data Preparation.....	20
3.4.	Frame Field Learning.....	23
3.5.	Polygonization Method.....	27
3.6.	Model Evaluation.....	28
3.7.	Experimental Setup.....	29
3.8.	Summary.....	29
4.	RESULTS AND DISCUSSION.....	30
4.1.	Quantitative analysis.....	30
4.2.	Qualitative analysis.....	31
4.3.	Over and Under-Segmentation.....	34
4.4.	Comparison of frame field results with standard segmentation model.....	36
4.5.	Testing transferability of the model.....	38
5.	CONCLUSION.....	40
5.1.	Reflection to research objectives and questions.....	41
6.	RECOMMENDATION.....	43

# LIST OF FIGURES

---

Figure 1: The concept of Fit-For-Purpose Land Administration and associated frameworks (Enemark et al., 2016).....	8
Figure 2: Conceptual framework of the research.....	12
Figure 3: Concept of the cadaster (FIG, 1995).....	14
Figure 4: Basic architecture of U-Net [adapted from Ronneberger et al., (2015)].....	17
Figure 5: General workflow of the study.....	18
Figure 6: Geographical location of the study area .....	19
Figure 7: A closer look at the agricultural fields in the study area.....	20
Figure 8: Snapshot of aerial (a) and UAV orthophoto (b) with corresponding cadastral boundaries (c).....	21
Figure 9: Dataset for the model is prepared in areas where UAV and aerial images overlapped (a), separate tiles for training, validation, and test sets (b) and dimension of a sample patch (c).....	22
Figure 10: Extended architecture of investigated frame field learning model [adapted from Girard et al. (2020b)].....	24
Figure 11: Detailed learning process of the frame field learning model.....	24
Figure 12: Post-processing and polygonization workflow [adapted from Girard et al. (2020)] .....	27
Figure 13: Pixel-wise evaluation using intersection over union metric.....	28
Figure 14: PoLiS distance indicated in black line between extracted(orange) and reference polygon vertices (blue). (Avbelj et al., 2015).....	29
Figure 15: reference data overlaid on the UAV orthophoto (a) predicted polygons on the UAV orthophoto (b) predicted polygons on aerial orthophoto (c).....	33
Figure 16: Over and under-segmentation cases: reference data overlaid on UAV orthophoto (black) prediction on UAV orthophoto (blue) prediction on aerial orthophoto (green) .....	34
Figure 17: Prediction on invisible boundaries: raw image (a) perceived holding rights indicated by the white arrows (b) predicted boundary (c).....	35
Figure 18: Detected boundary changes on the UAV orthophoto.....	35
Figure 19: Comparative results: Raw image (a), U-NetResNet101: trained with frame field (b) U-NetResNet101: trained without frame field (c).....	37
Figure 20: Results obtained using aerial image dataset on different geographical area. Reference data overlaid on aerial orthophoto (green) and model predictions on aerial orthophoto (yellow) .....	39

## LIST OF TABLES

---

Table 1: Summary of the final dataset used for model training and evaluation .....	22
Table 2: Training parameters of frame field learning model .....	30
Table 3: Summary of frame field learning model prediction results .....	30
Table 4: Quantitative results obtained on UAV and aerial test datasets.....	31
Table 5: Comparative test result of frame field and standard segmentation model .....	36

## ACRONYMS

---

ACM	Active Contour Model
BCE	Binary Cross Entropy
CNN	Convolutional Neural Network
DL	Deep Learning
ELU	Exponential Linear Unit
FCN	Fully Convolutional Network
FCN	Fully Convolutional Network with Dilated Kernels
FIG	International Federation of Surveyors
FFP-LA	Fit-For-Purpose Land Administration
GCP	Ground Control Points
IoU	Intersection over Union
ODM	Open Drone Map
PoLiS	Polygon and Line segments Similarity
ReLU	Rectified Linear Unit
ResNet	Residual Neural Network
SLLC	Second-Level Land Certification
UAV	Unmanned Aerial Vehicle
VHR	Very High Resolution

# 1. INTRODUCTION

## 1.1. Background

Land is the ultimate resource in which life cannot be supported without it. Thus, effective land administration is indispensable to ensure protection and proper use (Hull et al., 2020). According to UNECE (1996), land administration is defined as "the process of determining, recording and disseminating information about ownership, value, and use of land and its associated resources." It is imperative that land administration, as a system, meets the need of people-to-land relationships, provides tenure security for all, and sustainably manages natural resources (Enemark et al., 2014). Tenure security is an essential part of land administration function to reduce poverty and ensure food security by promoting investments on land (FAO, 2002). De Soto (2000) has re-emphasized the critical role secured property rights play in economic development. The United Nations Sustainable Development Goal, particularly indicator 1.4.2, promotes tenure security for all, special attention given to the poor and vulnerable groups (UN-HABITAT, 2018). But the reality on the ground tells a different story. It is roughly estimated that around 70% of the population in developing countries do not have access to a formal registration system to secure their land right (GLTN, 2012). In Sub-Saharan Africa, only 10 % of rural areas are legally registered; most of the others are undocumented or informally administered and thus prone to land grabbing (Byamugisha, 2013).

Ethiopia, a country located in Sub-Saharan Africa, is undertaking a massive land registration program to improve tenure security of smallholder farmers. The country has more than 50 million parcels of rural landholdings that need to be registered (Shibeshi Gebeyehu, 2011). In the late '90s, the government of Ethiopia has launched a large-scale land registration program called first-level land certification (FLLC). Individual plots are measured by using traditional tools such as rope. Although FLLC has achieved success in scale and outreach, the process only captured textual information of the parcel and tenure details. The limitation of FLLC is that it does not provide a cadastral map or sketch of the land (Ayano, 2018). Later in 2013, a new initiative known as second-level land certification (SLLC) was launched to register rural lands in areas where FLLC has been completed. SLLC is integrated with geoinformation technology to improve the land registration system (Cochrane and Hadis, 2019). Aerial photographs, satellite imagery, and GPS instruments are used for accurately measuring the location & size of plots on the field (Bezu and Holden, 2014). The benefits of SLLC are twofold. First, it ensures the detail of the landholders (textual information) and the boundary of parcels (spatial data) are verified and documented. Second, when subsequent transactions are recorded, landholders are provided with updated land certificates, ensuring that the register of interests in land is accurate and up to date (DAI, 2020). So far, the government has managed to cover one-third of the total parcels that have to be registered (Zain, 2021).

The fact that conventional land administration system has often been expensive and does not consider local conditions, its implementation, especially in developing countries, has shown little success (Van Asperen, 2014). With the current rate of land tenure recording, it would likely take centuries to achieve adequate land registration coverage (Zevenbergen et al., 2013). Notable organizations such as the International Federation of Surveyors (FIG) and Global Land Tool Network (GLTN) continuously advocated for the recognition of the continuum of rights and use of pro-poor recordation tools to

expedite land rights mapping (Hendriks et al., 2019). To that effect, Enemark et al. (2014) have introduced Fit-for-Purpose Land Administration (FFP-LA), an affordable, flexible & inclusive approach that mainly focuses on the purpose of the system. The most crucial concept in FFP-LA is the flexibility of the spatial framework (Figure 1), which includes (1) evolving through alternative data acquisition techniques such as satellite, aerial, or UAV images rather than deploying the traditional higher accuracy ground survey methods (2) use of general boundaries instead of placing expensive monuments to fix the extent of cadastral objects and (3) depending on the area, FFP-LA allows adoption of automatic feature extraction techniques for cadastral boundary extraction (Enemark and McLaren, 2018).

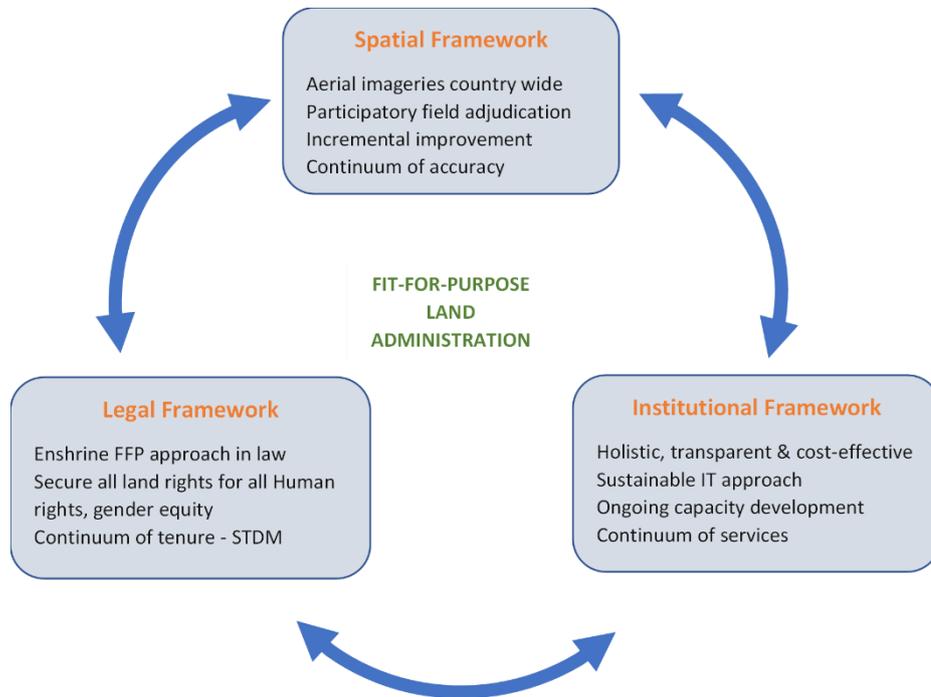


Figure 1: The concept of Fit-For-Purpose Land Administration and associated frameworks (Enemark et al., 2016)

The broader adoption of FFP-LA methodology has opened a new horizon for cutting-edge innovative technology, such as the use of unmanned aerial vehicle (UAV) and machine learning techniques to accelerate the cadastral mapping process (Kelm et al., 2021). UAVs are becoming increasingly affordable, reliable, easy to use and emerging as a data collection option for mapping applications (Colomina and Molina, 2014). It has been studied in depth by Stöcker et al. (2018) whether the use of UAVs in land administration can help meet the needs of developing countries. A thorough need assessment was carried out coupled with different UAV flight configurations in Rwanda to understand the usability. The authors concluded that UAV-based data collection sufficiently met the needs assessed beforehand. Horizon 2020, a European Union program, inspired by FFP principle have launched “its4land” project to deliver a suit of innovative solutions for mapping land rights mainly in sub-Saharan countries. The main deliverables of the project are (1) SmartSkeMa: a sketch based participatory mapping tool, that relies on indigenous knowledge of local communities to verify their own land rights (2) a UAV solution: which has introduced a complete UAV workflow starting from choosing the right type of UAV vehicle for the purpose of cadastral mapping to data collection and quality control steps (3) machine learning driven automatic cadastral boundary extraction. It consists of image pre-processing, segmentation, line extraction, contour

generation, and interactive post-processing steps. The workflow has been tested and proved to be valid on UAV, aircraft, and satellite RGB images. (4) A geo cloud service is now available through the Publish and Share (PaS) platform. The primary aim of PaS is to provide adaptable spatial reference system for land administration projects (Koeva et al., 2021).

Remote sensing is the science of acquiring information about an object without having to physically interact with it (V.A. Tolpekin and A. Stein, 2013). Trials have been done in the past to explore the feasibility of using high-resolution remote sensing images and feature extraction algorithms for cadastral boundary extraction (Bennett et al., 2020). Through (semi) automated boundary delineation, parcel-level information can be collected automatically in a scalable manner, thereby cutting down on costs and time (Aung et al., 2020). Kohli et al. (2017) carried out a visual comparison of cadastral boundaries in seven case study countries to investigate whether the boundaries are visible on corresponding high-resolution satellite images as a requirement for subsequent feature extraction tasks. The quantitative analysis revealed that the results range from 0%, where the topography is rugged and densely forested, to 71% in smallholder farms and planned urban settlements. This study revealed the potential of remote sensing-based methods for cost-effective and rapid cadastral mapping solutions. The study conducted by (Wassie et al., 2018) applied a semi-automatic feature extraction method to detect cadastral boundaries in a rural part of Ethiopia using the Mean Shift Segmentation algorithm on top of the WorldView-2 satellite imagery. The study used the buffer overlay method to evaluate the extracted results. They achieved 82.8%, 34.3% & 32.9% in terms of completeness, correctness, and quality, respectively. The algorithm gained good performance on flat landscapes and barely vegetated areas.

The early days of field boundary detection techniques typically rely on edge detection algorithms (Meyer et al., 2020); a primary hypothesis is that transition between fields exhibit an abrupt change of pixel values (Rydberg and Borgefors, 2001). There are many types of edge detection algorithm used in image analysis. Robert, Sobel, Prewitt, Laplacian of Gaussian (LoG) and Canny Edge Detection algorithms are well known for boundary extraction tasks (Muthukrishnan and Radha, 2011). The study by (Singh et al., 2015) showed how a modified version of Sobel Operator could extract parcel boundaries using 5 m. resolution Rapid Eye imagery. The traditional edge detection algorithms are unfavorable for satellite images because not all objects have sharp edges. They modified the existing operator to tackle this problem, by introducing a 5x5 mask to detect gradients in all directions, thus enhancing the detection result. Even though such techniques considerably reduce the manual digitization works, the detected edges are not regularized even some boundaries are left undetected, reducing the quality of extracted boundaries.

Improvements in spatial resolution of satellite and aerial imagery and ongoing developments in computer vision made it possible to detect objects with increasing accuracy and (semi) automatic procedure (Avbelj et al., 2015). Recent advances in deep learning (DL) methods that systematically learn multi-level features have given geospatial application domains a new outlook for extracting information from very high resolution (VHR) remote sensing images (Sirko et al., 2021). Neuronal networks are capable of learning any function, even high dimensional and nonconvex problems. By semantically segmenting pixels based on whether they belong to a class of interest or not, this method offers better granularity and could be used directly without edge detection stage (Masoud et al., 2019). With Convolutional Neural Networks (CNN), (Crommelinck et al., 2019) demonstrated how boundary line classification can be conducted without needing hand crafted features, achieving 71% accuracy. In addition, Xia et al. (2019) used Fully Convolutional Networks (FCN) to detect cadastral boundaries in urban and peri-urban environments in Rwanda using UAV orthophoto. Overall, FCN achieved significant performance compared to other state-

of-art methods such as Multi Resolution Segmentation (MRS) and gPb-based contour detection methods. Reported result in precision, recall, and F-score, are 0.79, 0.37, and 0.50, respectively. A recent study by Fetai et al. (2021) investigated detection of visible boundaries from UAV image in rural Slovenia using U-Net architecture and compared the results with ENVI DL, a commercial software module. The authors achieved a balanced result from U-Net model with recall and precision of 0.65 and 0.41 respectively compared to the commercial software-based ENVI DL module which achieved 0.84 and 0.35 recall and precision. Garcia-Pedrero et al. (2019) trained UNET model to automatically delineate boundaries of agricultural fields using aerial orthophoto & publicly available Land Parcel Identification System (LPIS) dataset. The result obtained from UNET has outperformed gPb-UCM method with 89% overall accuracy. The quality of predicted boundaries are evaluated using boundary displacement index (BDE) by measuring distance between boundary pixels in the segmented image compared with reference data boundary pixels. They achieved BDE index of 14.64 for gPb-UCM and 5.53 for UNET. One of the major issues encountered in their study was that what is seen in the image is different from what is perceived as a parcel boundary. This issue stands in the way of the automatic extraction process.

## 1.2. Research problem

Many geographical information system applications, including cadastral mapping, require vector representation of spatial objects. Although DL networks such as CNN and FCN have achieved remarkable result in segmenting remote sensing images, the final map of the segmentation output is usually in a raster format (Zorzi et al., 2021). As a result, the segmentation output cannot be directly consumed for mapping applications. Generally, previous works on boundary extraction use two different approaches to extract vector polygons from the segmented image (Girard et al., 2020a). The first method is to convert the probability map produced by the neural network into a vector map followed by simplification algorithms, such as Ramer-Douglas-Peucker, to eliminate unnecessary vertices. Such techniques suffer from overly smoothed and irregular outlines. As a result, it imposes enormous post-processing task. The second method directly predicts object outlines from the input image in an end-to-end fashion. Zhao et al. (2021) introduced an improved method that can directly predict the outline of buildings in vector format by integrating CNN and Recurrent Neural Network (RNN) architectures. The CNN learns to find the location of vertices of the target object and RNN connects them together to create the outline. To accommodate the need for refined and regularized building outlines, modifications have been made to various network modules. The method was tested on high-resolution images. The evaluation result shows better COCO & PoLiS metric performance than state-of-art Mask-RCNN and PolyMapper. Persello et al. (2019) develop a strategy to delineate agricultural field boundaries using FCN in two different study areas. The initial image segmentation task was done using SegNet network. After segmentation, oriented watershed transform (OWT) is introduced to generate closed contours from the fragmented raster output. The final accurate field boundary is acquired after applying a grouping algorithm. The results are reported using precision-recall metric and they obtained F-score of 0.7 and 0.6 for two study areas.

There are limitations to the studies mentioned above, such as the fact that RNNs are computationally intensive to train, and these kinds of methods do not reuse the same boundary across adjacent parcels. Girard et al. (2020) proposed a novel building outline detection algorithm based on frame field learning that can directly output vector polygons to solve the above mentioned issues. Initially, UNET-16 and DeepLabV3, an encoder-decoder architecture, were used as a backbone to perform the segmentation task. A frame field is simultaneously learned to improve the segmentation quality and guide the subsequent polygonization process.

Many studies used state-of-art DL methods for cadastral boundary extraction focused for the most part on building outline extraction. But only a few studies have used neural networks, specifically FCNs, for agricultural field boundary delineation. It has not yet been reported in any previous studies that DL methods have been used to extract agricultural field boundaries in a vector polygon format. By taking advantage of new DL networks and high-resolution remote sensing images, this study proposed a method for extracting cadastral boundaries of agricultural fields in a vector format. Vector data are essential for mapping applications since they are the basis for land information systems.

### **1.3. Research objective and questions**

#### **1.3.1. General objective**

The main objective of this study is to design a method that can extract cadastral boundaries from very high-resolution remote sensing images in a vector polygon format.

#### **1.3.2. Sub-objective and questions**

1. To investigate deep learning methods used for cadastral boundary extraction
  - 1.1. What methods exist for cadastral boundary extraction?
2. To prepare data for cadastral boundary extraction
  - 2.1. What are the steps needed to post-process UAV images?
  - 2.2. How to design & prepare input dataset for the model?
3. To design a deep learning method used for boundary extraction
  - 3.1. Which FCN based network architecture is suitable to use as backbone?
  - 3.2. How to get optimum hyperparameters of the model?
  - 3.3. What methods are used to extract regularized vector polygons?
4. To evaluate the performance of the method
  - 4.1. What are the standard metrics used for evaluating predicted polygons?
  - 4.2. What is the performance of the model on UAV and aerial orthophoto?

### 1.4. Conceptual framework

The main concepts explored in this thesis are shown in Figure 2. As previously mentioned, there have been limited attempts to employ automatic feature extraction especially deep learning methods in combination with VHR remote sensing images for the purpose of cadastral mapping. In this research, inspired by FFP spatial framework principle, the potential of feature extraction algorithm is explored and applied on remote sensing images to test their usability for cadastral mapping. The spatial framework of the FFP-LA promotes remotely sensed images as an alternative method to high precision ground surveys. In addition, general boundary principle is used to delineate boundary of parcels. Feature extraction methods such as DL methods can be applied to the components of the spatial framework to facilitate the cadastral mapping process. The goal of this innovative method is to support tenure security.

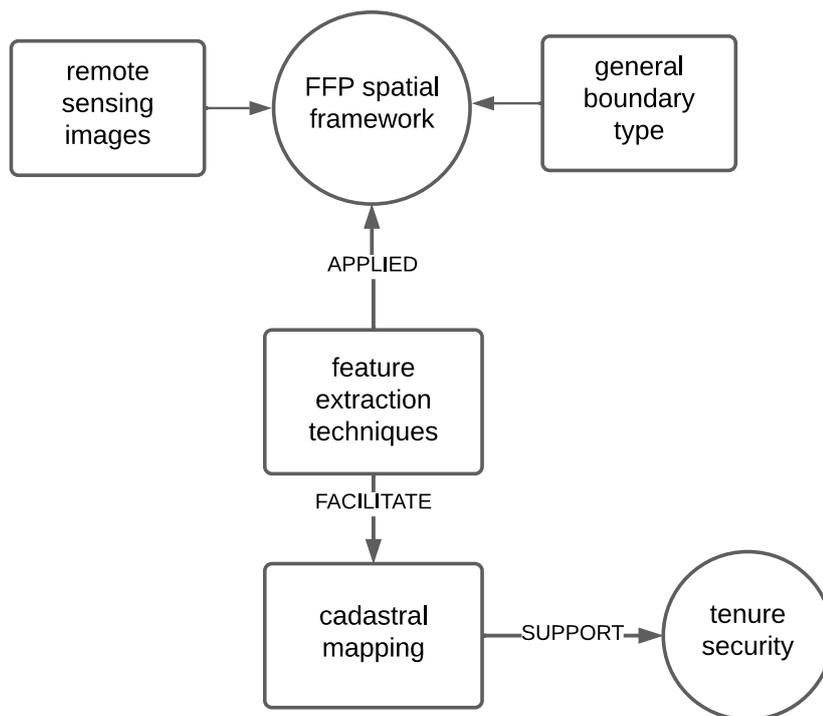


Figure 2: Conceptual framework of the research

### 1.5. Ethical considerations

The datasets collected for the purpose of this thesis has been used in compliance with the EU General Data Protection Regulation (GDPR). Personal information of the land holder is removed from the reference cadastral boundaries before doing further processing on the data.

## 1.6. Structure of the thesis

This thesis shows the usability of deep learning models to extract cadastral boundary extraction from VHR remote sensing images in a vector polygon format. It is organized in six chapters.

### **Chapter one: Introduction**

In this chapter, background and justification of the research problem is discussed in detail. Research objective and questions are formulated to address the research problem. Furthermore, the conceptual framework provides summary of the main concepts that will be discovered in this thesis.

### **Chapter two: Literature Review**

An extensive literature review of different journal articles, books, and academic outputs about automatic cadastral boundary extraction using FCN based deep learning networks are presented in this chapter.

### **Chapter three: Research Methodology**

In this chapter the proposed frame field learning method is discuss in detail. A description of the datasets used in this study and how they were prepared is also provided. Furthermore, the proposed pixel and polygon level metrics to evaluate model results are introduced

### **Chapter four: Results and Discussion**

The results of model experiments are presented in this chapter along with discussion of results. This is the order in which the results are presented. Comparison of polygon predictions based on UAV and aerial orthophoto, comparison of frame field learning with standard segmentation model and finally testing transferability of the model.

### **Chapter five: Recommendations**

Based on the work of this research, recommendations are provided for future works aimed at cadastral boundary extraction in vector polygon format.

## 1.7. Summary

In this chapter, the societal problem related to tenure security is briefly discussed that leads to the research problem which is cadastral boundary extraction in vector polygon format. Research gaps are outlined in this section. The research objectives and questions are formulated in context with the background of the core aspects of the research which helps to identify appropriate research methodology that can solve the research problem. In summary, this research aims to take advantage of new deep learning networks to extract cadastral boundaries from VHR remote sensing images in a vector polygon format.

## 2. LITERATURE REVIEW

### 2.1. Cadastral boundaries & mapping methods

FIG in its statement described the cadastre as "a parcel-based and up-to-date land information system containing a record of interests in land (e.g., rights, restrictions and responsibilities). It usually includes a geometric description of land parcels linked to other records describing the nature of the interests, and ownership or control of those interests, and often the value of the parcel and its improvements." (FIG, 1995). Parcel objects are the heart of a cadastral system. In real life, a parcel of land is contiguous piece of land that encompass unique and homogeneous legal and use interests (Henssen J.L.G, 1995).

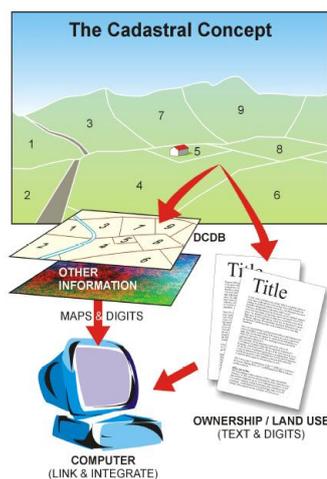


Figure 3: Concept of the cadaster (FIG, 1995)

Cadastral maps serve as a reference document that describes geographical location and extent of a property (Dale, 1977). Depending on how data is collected, cadastral maps can be created in either direct or indirect ways. In a direct surveying, an instrument such as GNSS, Electronic Distance Meters (EDM), theodolite or measuring tape is applied directly to the unknown quantity, and the value of the quantity is observed. Surveyors have long used this technique to determine the boundary of a property with high precision. The second method, Indirect surveying, is carried out using data gathered from remote sources, such as satellites, aircraft, or UAVs. The boundaries are then extracted by digitizing manually on the image (Corlazzoli and Fernandez, 2004).

A cadastral boundary is classified as fixed or general based on how it is defined. Fixed boundaries are delineated more precisely and are often demarcated with permanent constructions or monuments on the ground. On the other hand, general boundary is one in which the precise location is left undescribed, but the boundary is usually, but not always, indicated by a physical feature. (Nichols, 1993). Some examples of physical features include fence, hedge, waterway, and tracks (Luo et al., 2017).

Depending on the boundary concept one chooses different techniques available to acquire cadastral data. Aerial photography would be most economically viable if a large area is being considered for systematic adjudication (Tuladhar, 1996). For a long time, orthorectified aerial photographs has been essential for the creation of cadastral maps. Aerial photographs meet tolerable measurement requirements to produce

cadastral maps at rural and urban scale (IAAO, 2004). Even more, the accuracy of UAV-based mapping can be measured down to centimetres, which helps identify smaller objects (Eisenbeiss, 2011). This study relies on general boundary principle to extract cadastral boundaries of agricultural fields from VHR aerial images using fully convolutional networks and frame field learning.

## **2.2. Image-based automatic boundary extraction**

Establishing cadastral maps based on traditional ground surveying (direct methods) could be a lengthy process, expensive, and require a lot of manual work (Enemark et al., 2016). Due to their convenience and efficiency, photogrammetric techniques offer a simple and comprehensive solutions to various mapping problems mentioned earlier (Blachut, 1976). Through these techniques, one can extract boundary of a property, with a certain tolerance level that would be obtained using a conventional survey instrument without engaging laborious effort on the field. As remote sensing technology has become ubiquitous, it is easier and more affordable to obtain remote sensing images for the purpose of indirect surveying.

It is often the case that orthorectified aerial images were digitized manually by human operators to extract visible features. However, recent developments in computer vision and digital image processing have made it possible for these extraction processes to be delivered in automatic or semi-automatic way (Eker and Seker, 2008). Extracting cadastral boundary from remotely sensed image is treated as segmentation problem which is a recurring issue in remote sensing image classification (Oliveira et al., 2017). Convolutional neural networks (CNN) have gained considerable success due to their capability to extract high-level features without requiring human expertise in feature engineering. An in-depth technical explanation of deep learning methods for remote sensing data is given by Zhang et al. (2016). CNN is a neural network architecture that incorporates several feature extraction levels. At each level, convolution, non-linearity activation and pooling layers are executed. The feature extraction level is followed by one or more dense layer and final classification layer. CNNs are characterized by their deep structure which provides the model with the capability of learning salient features and map these features into representations that may aid in improving the performance of successive classifiers.

Due to the limitation that CNNs can only provide patch-based predictions, meaning that the prediction would only contain one class per patch, there was a need for semantic segmentation, in which every pixel is assigned to a class. Fully convolutional network (Shelhamer et al., 2017) was introduced as a modification to the standard CNN architecture to resolve this issue. There are many varieties of FCN based neural network architectures. However, the components of these networks are mostly similar.

## **2.3. Fully Convolutional Neural Networks**

A fundamental concept in remote sensing is pixel-based classification which involves assigning a semantic class to every pixel of the input image. Fully Convolutional Networks are primarily used for per pixel classification (He et al., 2019). FCNs use only convolution, pooling, and up-sampling as locally linked layer. The network can be trained faster by using fewer parameters since fully connected nodes are replaced by up sampling (deconvolution) layer. Deconvolution sometimes known as transposed convolution is a reverse operation of the traditional convolution to recover downsampled feature maps back to the original resolution to enable pixel-based classification. Deconvolution first upsamples the feature map by a factor of the stride and padding value, followed by a convolution operation on the upsampled feature map by memorizing the pooling indices used to downsample the input map.

Furthermore, since all connections are local, FCNs can handle a wide range of image sizes (Shelhamer et al., 2017).

FCN with dilated kernels (FCNDK) has been found useful for tasks that needs larger spatial context. By using dilated convolution, features can be captured with a larger receptive field without being downsampled (Persello and Stein, 2017). Previous works by Musyoka (2018)(Aung et al., 2020), Masoud et al. (2019), Crommelinck et al. (2019) and Xia et al. (2019) has proven FCN based networks to be effective in extracting cadastral boundaries from remotely sensed images. Detailed cases and results achieved by FCN based networks can be found in chapter 1.

## 2.4. UNET

U-Net is an encoder-decoder architecture developed by Ronneberger et al. (2015) originally for biomedical image segmentation. It was brought by modifying and extending the existing FCN architecture in such a way that it works with very few training images and delivers more accurate segmentation output. The encoder path performs two 3x3 convolution, ReLU activation, and a 2x2 max pooling to learns high-level features from the input image. The decoder path performs a 3x3 convolution over the concatenated image and feature map from the encoder path to recovers spatial information lost during down-sampling process. Skip connections solve the degradation problem and provide localized information by cropping the full size image from encoder path and concatenating them directly to the decoder path (Drozdzal et al., 2016). The final layer holds a classification map with the desired number of classes.

Besides its original application for biomedical images, UNET has been widely used by several studies in the geospatial domain to extract information from remote sensing data. A study by Wierzbicki et al. (2021) has implemented UNET to delineate building outlines by fusing VHR aerial orthoimages and LiDAR data for the purpose of cadastre modernization in Poland. The result is reported based on precision-recall framework. They achieved overall accuracy of 89.5% which indicates the model is well suited for use in building extraction. Precision of 78.5% and 80.7% recall. According to the study, deep learning methods hold promise for updating cadastres in the future. (Wang et al., 2022) has experimented FracTAL-ResUNET, an attention-based UNET, to delineate field boundaries in France and India using two different imageries from SPOT and PlanetScope. They achieved mean IoU of 0.86 and 0.72 respectively on both images. The results suggest that the method provides a scalable approach to identifying field boundaries in regions where there are no cadastral datasets are available. Aung et al. (2020) experimented UNET to extract agricultural fields in France by taking temporal information into consideration. This approach allows them to achieve 83% accuracy compared to other configurations of UNET. In addition, Fetai et al. (2021) and Garcia-Pedrero et al. (2019) experimented using UNET and concluded that they could accurately extract cadastral boundaries from high-resolution aerial images with little post-processing required. Both studies achieved overall accuracy of above 89%.

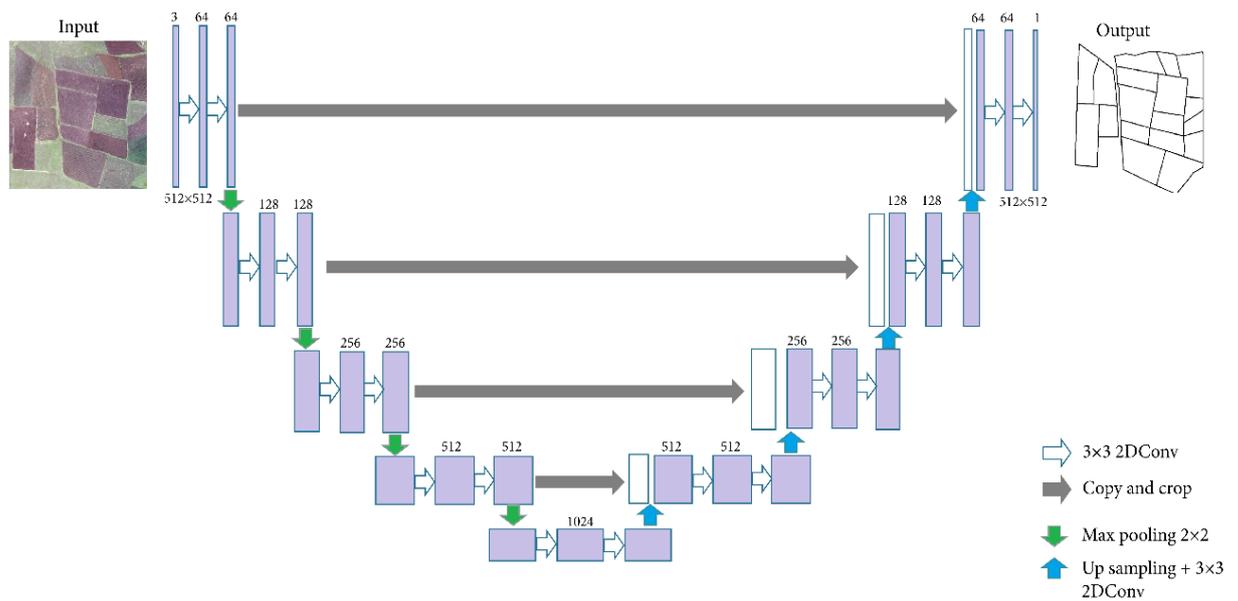


Figure 4: Basic architecture of U-Net [adapted from Ronneberger et al., (2015)]

## 2.5. SegNet

SegNet is another auto-encoder-inspired network which consists of an encoder-decoder like architecture followed by soft-max classification layer. Despite a few notable differences, the architecture is similar in many ways with U-Net. The decoder part of SegNet memorizes pooling indices to upsample the lower resolution feature maps created during pooling operation on the corresponding encoder path. Additionally, the convolution layers in the decoder network are identical to those in the VGG-16 network. With SegNet, a larger context is considered with increasing depth, improving pixel labelling accuracy (Badrinarayanan et al., 2015).

Persello et al. (2019) used SegNet to delineate boundary of agricultural parcels by utilizing Worldview satellite imagery. The experimental study was carried out in Nigeria and Mali where small holder agriculture is the norm. The results from SegNet have outperformed alternative methods by achieving F1 score of 0.7 and 0.6 for both case studies, respectively. The agricultural fields were smaller in size, crops are variable within one field and spectral characteristics is very dynamic in space and time. Such cases become a major obstacle and compromise the quality of boundary detection in small holder agriculture environment.

## 2.6. Summary

In this chapter, core concepts related to the spatial framework of FFP-LA, more specifically, cadastral boundary types and image-based mapping methods, are explained in detail. Furthermore, previous works on cadastral boundary extraction using deep learning techniques were thoroughly reviewed. Several studies have shown that FCN-based deep learning networks are more effective for detecting cadastral boundaries. FCNDKs, SegNet and UNET were the focus of the review. In the experimental setup phase, all these FCN based network architectures will be tested and the best performing architecture will be used as a backbone for the main vector polygon boundary extraction model.

### 3. RESEARCH METHODOLOGY

#### 3.1. General workflow

The research methodology of this study is comprised of four main parts. First, a systematic literature review is conducted to identify the most relevant neural network architecture for the task of cadastral boundary extraction. Additionally, the comparison metrics at both the pixel and polygon level are examined for use in model evaluation. The second step is data preparation. The UAV image is orthorectified using GCPs to generate true orthophoto. To train the DL model, 1024x1024 patches are prepared from both UAV and aircraft orthophoto and the cadastral boundaries using image tiling technique. On the third step, several experiments were conducted to find optimum hyperparameters of the model. With these parameters the frame field model is trained and the subsequent polygonization method follows to extract polygons in a vector format. Finally, the model results are evaluated using the metrics mentioned earlier. Polygons predicted using UAV and aerial orthophoto are compared against each other.

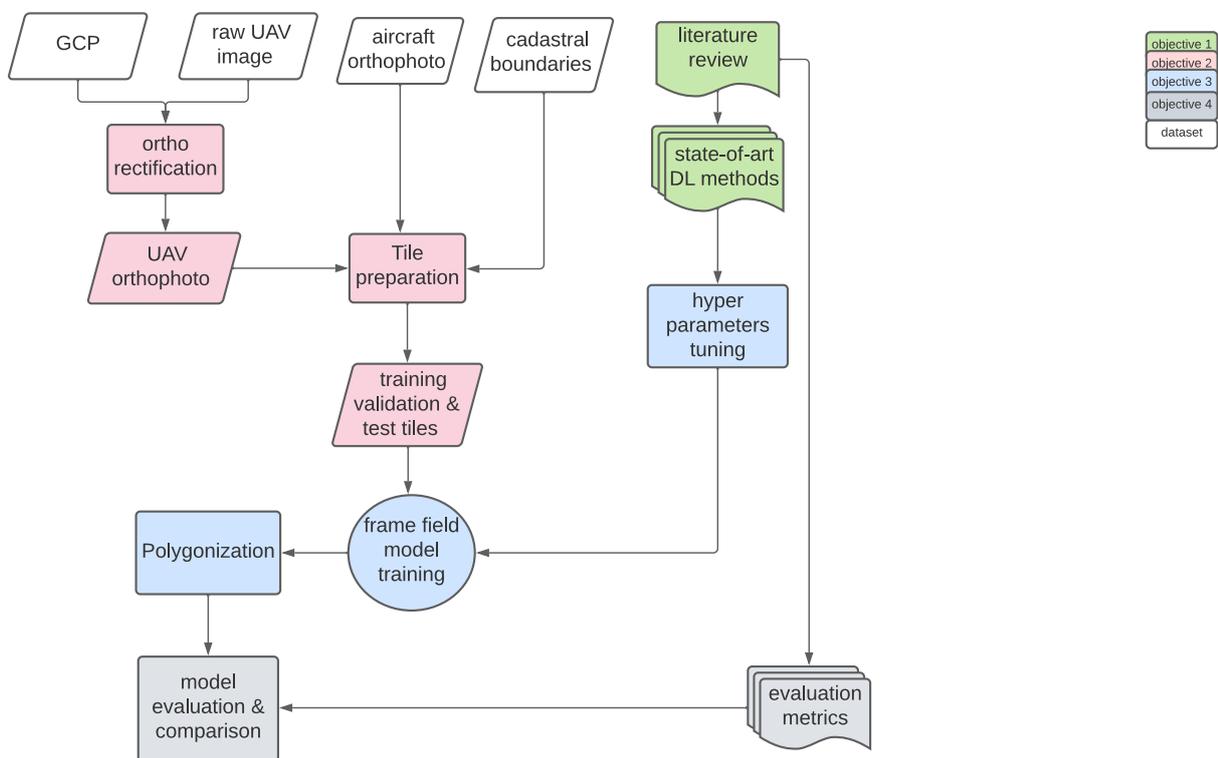


Figure 5: General workflow of the study

### 3.2. Study area & datasets

#### 3.2.1. Location

The study area is located in North Shoa Zone, central part of Ethiopia. The selection is based on the availability of high-resolution images, full cadastral coverage, and suitability of the topography for automatic feature extraction. The study area covers one kebele which is the smallest administrative unit or ward in Ethiopia. Small holder farming is the most dominant activity in the area.

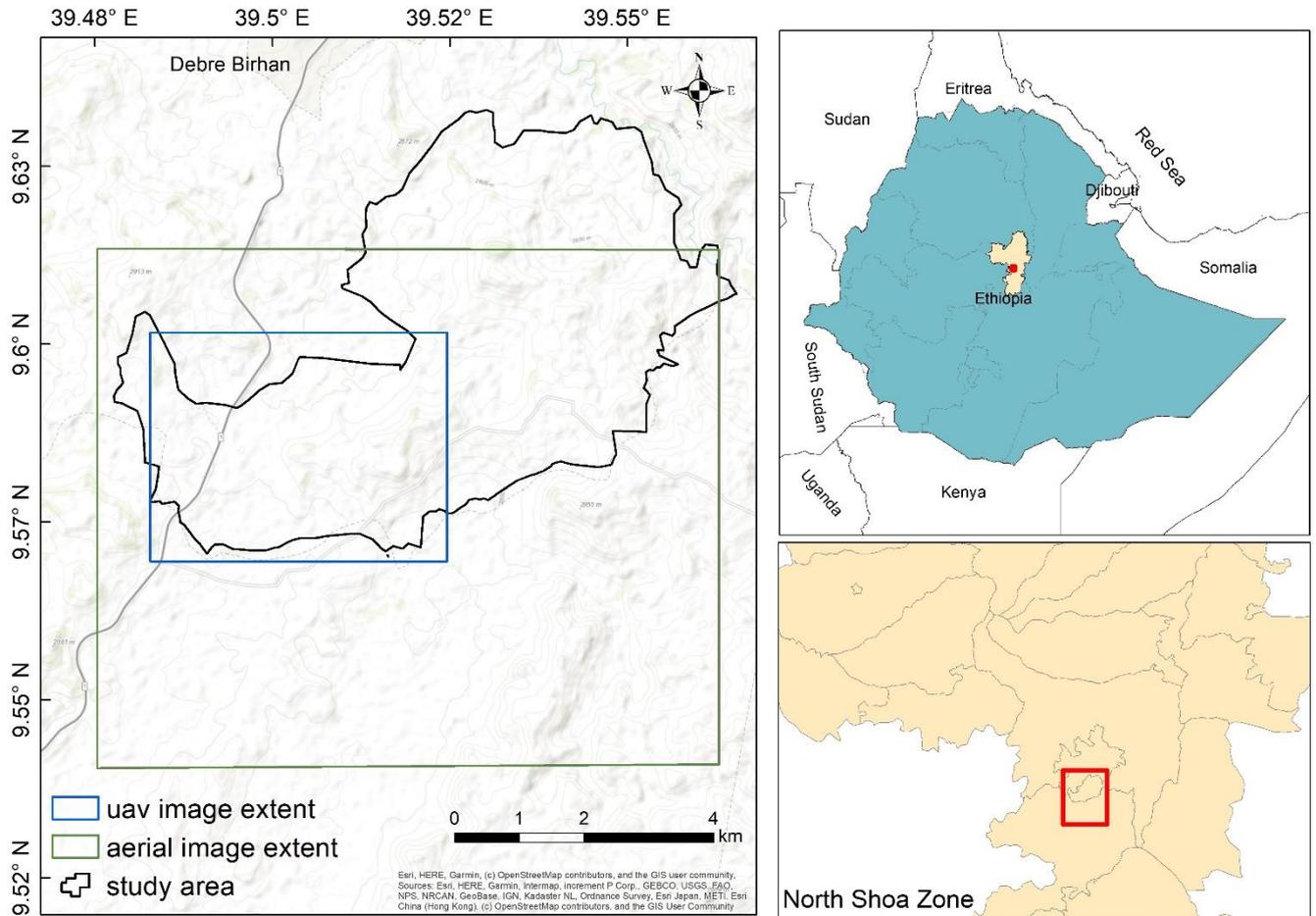


Figure 6: Geographical location of the study area

In 2013 Ethiopia has launched second level land registration in rural part of the country. The implementation of SLLC is based on the principles of FFP-LA. Country wide aerial orthophoto have been used to demarcate the boundaries of agricultural parcels in a participatory mapping approach (Milindi Rugema et al., 2020)

In rural parts of Ethiopia, it is common to observe hedges marking the border between two agricultural fields. The country follows general boundary principle and use of large scale orthophoto to delineate agricultural parcels in rural areas (Ministry of Agriculture, 2016). As shown in figure 7 below, hedges indicated by white arrow are remained untouched for boundary reference despite the crop was fully harvested.



Figure 7: A closer look at the agricultural fields in the study area

### 3.3. Data Preparation

As this study follows the principles of fit-for-purpose land administration, the data preparation pipeline was fully implemented using free and open-source software (FOSS) platforms aiming to provide low-cost and scalable alternatives. Imagery and reference cadastral boundary data is provided by Ministry of Agriculture.

#### 3.3.1. Imagery

For this study, orthophotos based on images captured from two different aerial platforms are used to train the deep learning model. The UAV images were acquired in 2019 with 0.11 m spatial resolution and covers 9 sq. km. The orthophoto based on the UAV imagery was created within the work of this research. The digital aircraft images were taken in 2017 and the already produced orthophoto with a spatial resolution of 0.29 m. It covers an area of 50 sq. km. Both UAV and digital aircraft data acquisitions were taken in the visible spectrum range (RGB channels). Throughout this study the aircraft orthophoto is referred as aerial orthophoto.

#### 3.3.2. Reference Data

Deep learning models require the use of large-scale labelled data sets (Zhang et al., 2016). Thus, the already demarcated agricultural parcel boundaries are used as a reference data. The cadastral boundaries are delineated from 0.29 m resolution orthophoto. in Ethiopia, dense agricultural areas are demarcated with 0.5 to 1m allowable error (Ministry of Agriculture, 2016). Therefore, a 0.5 m buffer around the boundary is added to allow certain tolerance when predicting the boundaries.



Figure 8: Snapshot of aerial (a) and UAV orthophoto (b) with corresponding cadastral boundaries (c)

### 3.3.3. Orthophoto generation

Orthophoto mosaic is created from the UAV's raw images with the help of open drone map (WebODM), an open-source command line utility for processing aerial imagery. WebODM has been found robust and relatively automated compared to other open-source counterparts (Vacca, 2019). Using the GCPs and intrinsic and extrinsic parameters of the camera, ODM automatically creates the UAV orthophoto.

### 3.3.4. Tile preparation

Using custom image size to train a deep learning model would typically cause memory bottleneck problem. This issue can be solved by using a technique called "tiling". The procedure involves cutting up the orthophoto and corresponding cadastral boundaries into smaller square patches. Afterward, depending on the batch size, each of these patches are analyzed individually to save time and computing resources. Solaris, an open-source python library, is used to perform the tiling operation. To reduce computational load on the server, the size of training and validation patch is set to 512\*512 pixels. In terms of area, each patch measures 1 hectare, which provides a good context for agricultural field. For a larger area prediction, the size of test patches was set to 1024x1024 pixels (Figure 9c)

### 3.3.5. Dataset split

In supervised learning, the goal is to build a model that performs well on previously unseen dataset. Guided by this principle, the whole dataset is systematically divided into training, validation, and test sets on areas where the UAV and aerial orthophotos are overlapped (Figure 9a). Patches that contain less than 20% reference data are removed. To choose optimal hyper parameters of the model, a validation set is randomly generated from the training dataset. To prevent bias in the accuracy assessment due to spatial autocorrelation, the test tiles are spatially disjointed from the training tiles (Figure 9b).

Dataset	Number of patches	Patch size	Proportion
Training	474	512	0.66
Validation	120	512	0.17
Test	61	1024	0.17

Table 1: Summary of the final dataset used for model training and evaluation

to avoid shortage of training data and model over-fitting, data augmentation technique is applied on the training dataset. The goal is to significantly increase the size of data available for training models by creating artificial dataset with custom transformations. Position and colour augmentation such as scaling, rotation, and colour jitter transformations are applied before feeding it to the network.

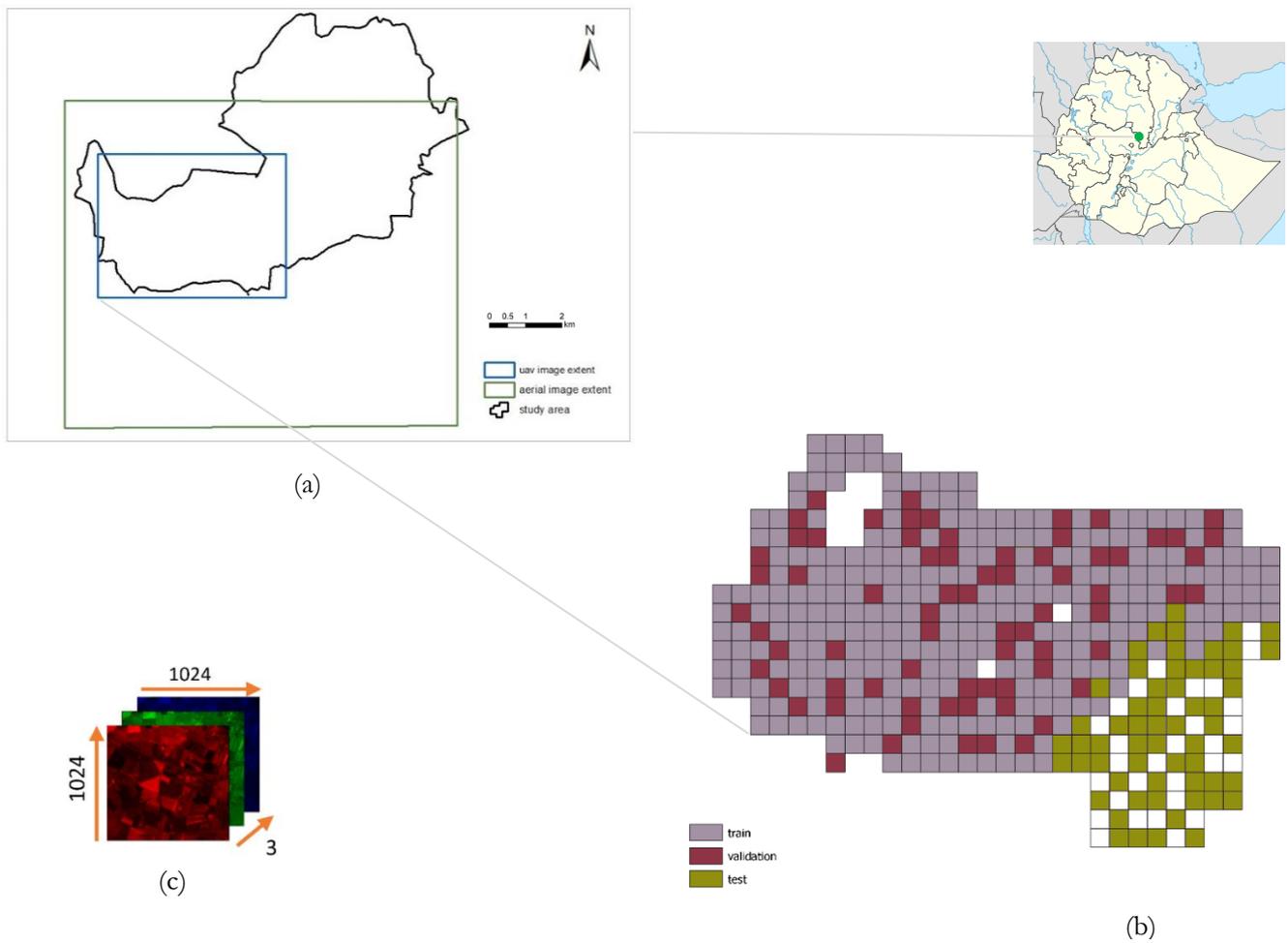


Figure 9: Dataset for the model is prepared in areas where UAV and aerial images overlapped (a), separate tiles for training, validation, and test sets (b) and dimension of a sample patch (c)

### 3.4. Frame Field Learning

Classification maps produced by a standard segmentation model such as UNET may have many ambiguous cases due pixels discontinuity (noise). This results in irregular edges and rounded corners. Consequently, the output cannot be used directly for mapping applications, therefore extensive post-processing is required. To overcome the above problem, Girard et al. (2020) proposed a novel building outline detection algorithm based on frame field learning that can directly output regularized vector polygons. FCN is used as a backbone for pixel-based image segmentation and frame fields are added to obtain regularized building polygons. Like many building structures, agricultural land parcels have reasonably regular (near straight) edges and corners. In this study, the frame field learning algorithm is, therefore, adapted for extraction of agricultural parcel boundaries in vector polygon format.

#### 3.4.1. Frame Field

As presented in the work of Diamanti et al. (2014), Frame Field is a 4-poly vector field that assigns two pairs of directional vectors for every point on a plane. The first two vectors are placed symmetrical of the other two. In other words, every point on image plane is linked to a set of vectors  $\{\mathbf{u}, -\mathbf{u}, \mathbf{v}, -\mathbf{v}\}$  which defines the frame field by two complex numbers  $\mathbf{u}, \mathbf{v} \in \mathbb{C}$ . Their polynomial representation is given in the following formula.

$$f(z) = (z^2 - u^2)(z^2 - v^2) = z^4 + c_2 z^2 + c_0 \quad (1)$$

The constants  $c_2$  and  $c_0$  in (2) uniquely identifies a frame  $\{\mathbf{u}, \mathbf{v}\}$  at every pixel position in the image (Bessmeltsev and Solomon, 2018). Recovering frame direction from the constants can be easily done as follows:

$$\begin{cases} c_0 = u^2 v^2 \\ c_2 = -(u^2 + v^2) \end{cases} \Leftrightarrow \begin{cases} u^2 = -\frac{1}{2} \left( c_2 + \sqrt{c_2^2 - 4c_0} \right) \\ v^2 = -\frac{1}{2} \left( c_2 - \sqrt{c_2^2 - 4c_0} \right) \end{cases} \quad (2)$$

The frame field learning training procedure is depicted in Figure 10 below. Given a 3-channel image as input, the model outputs a segmentation map and frame field. The segmentation map produces edge mask which corresponds to parcel boundaries ( $\hat{y}_{edge}$ ) and an interior mask corresponds to extent of the land parcel ( $\hat{y}_{int}$ ). The frame field produces two pairs of directional fields which corresponds to the two complex coefficients  $c_2, c_0 \in \mathbb{C}$ .

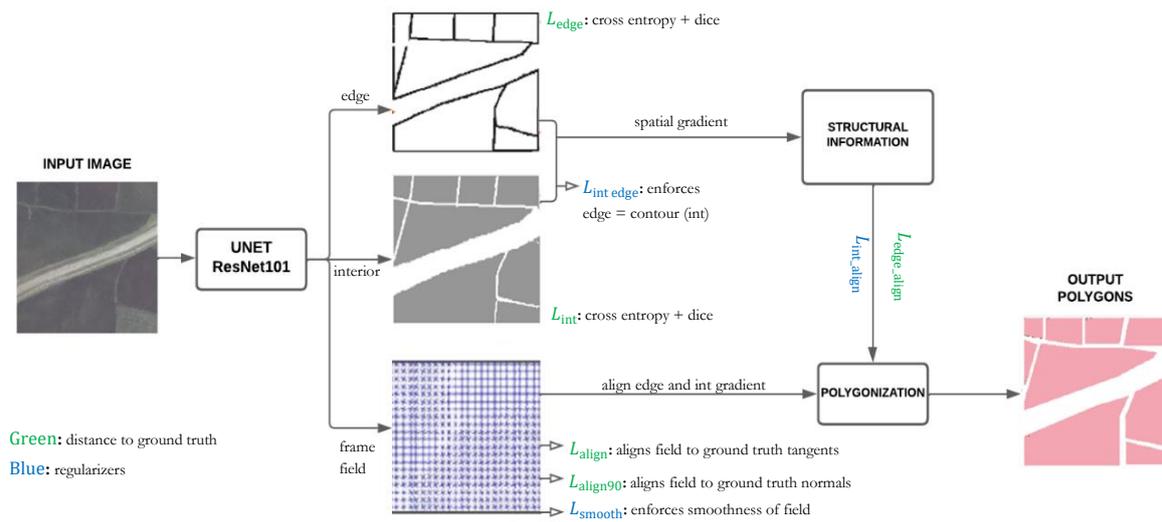


Figure 10: Extended architecture of investigated frame field learning model [adapted from Girard et al. (2020b)]

Along the edges of a parcel, at least one vector field is aligned with the polygon tangent direction. On the other hand, Poly-Vector fields are aligned to both tangent directions at polygon corners to capture bi-directionality (Girard et al., 2020b).

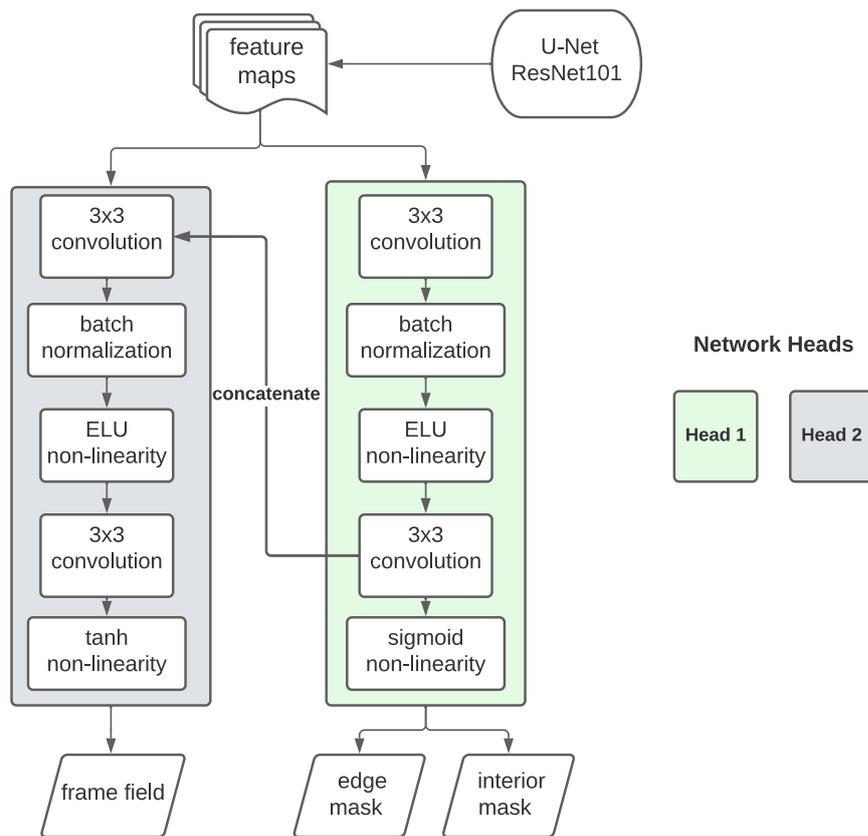


Figure 11: Detailed learning process of the frame field learning model

As shown in Figure 11, the proposed model generates three different outputs simultaneously. The first head of the network takes feature maps generated by the backbone as input and delivers edge and interior masks. The second network head outputs direction information in the form of a frame field by concatenating feature maps generated by the backbone plus segmentation maps created by the first network head.

### 3.4.2. Backbone

U-Net is chosen as a backbone to perform the segmentation task. The encoder part of the network is replaced by a pretrained ResNet101 encoder. This has two benefits. First, it speeds up training time. Second, the model will have a starting knowledge of the kind of features that need to be detected.

### 3.4.3. Losses

By aligning masks and the frame field to the reference data, the loss ensures that the outputs are smooth and consistent. During model training phase, various loss functions for different learning tasks has been implemented.

### 3.4.4. Segmentation Loss

Inherently the segmentation task is performed by a neural network specified in the backbone. A pretrained U-Net with ResNet101 encoder, is the preferred backbone to output the segmentation masks. A combination of Cross Entropy and Dice Loss, a typical solution to handle the issue of class imbalance (Sudre et al., 2017), has been used as a loss function. The combined loss is computed using the equation given below:

$$L_{int} = a \cdot L_{BCE}(\hat{y}_{int}, y_{int}) + (1 - a) \cdot L_{Dice}(\hat{y}_{int}, y_{int}) \quad (3)$$

$$L_{edge} = a \cdot L_{BCE}(\hat{y}_{edge}, y_{edge}) + (1 - a) \cdot L_{Dice}(\hat{y}_{edge}, y_{edge}) \quad (4)$$

Where  $L_{BCE}$  is the standard Cross Entropy loss and  $L_{Dice}$  is Dice loss. The hyperparameter is denoted by Alpha ( $a$ ). Both losses are applied simultaneously on edge ( $L_{edge}$ ) and interior ( $L_{int}$ ) masks generated by the backbone.

### 3.4.5. Frame Field Loss

Besides the segmentation masks generated at the previous step, the model laterally outputs directional information of each pixel in the form of frame field. The three losses  $L_{align}$ ,  $L_{align90}$ , and  $L_{smooth}$  are simultaneously learned when training the frame field.  $L_{align}$  ensures alignment between the frame and the tangent vector direction.  $L_{align90}$  constraints the frame field not to collapse into a line. Finally,  $L_{smooth}$  enables the network to produce a seamless frame field. The frame field loss is denoted by the following equation

$$L_{\text{align}} = \frac{1}{HW} \sum_{x \in I} \hat{y}_{\text{edge}}(x) f(e^{i\theta_{\tau}; C_0(x), C_2(x)})^2 \quad (5)$$

$$L_{\text{align } 90} = \frac{1}{HW} \sum_{x \in I} \hat{y}_{\text{edge}}(x) f(e^{i\theta_{\tau \perp}; C_0(x), C_2(x)})^2 \quad (6)$$

$$L_{\text{smooth}} = \frac{1}{HW} \sum_{x \in I} (\|\nabla C_0(x)\|^2 + \|\nabla C_2(x)\|^2) \quad (7)$$

### 3.4.6. Coupling Loss

As explained on the prior steps, throughout the frame field learning procedure, the model produces various outputs simultaneously. The purpose of coupling loss is to maintain consistency between these outputs to create strong correlation amongst each other.

$$L_{\text{int align}} = \frac{1}{HW} \sum_{x \in I} f(\nabla y_{\text{int}}(x); C_0(x), C_2(x))^2 \quad (8)$$

$$L_{\text{edge align}} = \frac{1}{HW} \sum_{x \in I} f(\nabla y_{\text{edge}}(x); C_0(x), C_2(x))^2 \quad (9)$$

$$L_{\text{int edge}} = \frac{1}{HW} \sum_{x \in I} \max(1 - y_{\text{int}}(x), \|\nabla y_{\text{int}}(x)\|_2) \cdot \left| \|\nabla y_{\text{int}}(x)\|_2 - y_{\text{edge}}(x) \right| \quad (10)$$

$L_{\text{int align}}$  and  $L_{\text{edge align}}$  ensures that the spatial gradient of edge and interior masks are aligned correctly with the frame field (equivalent to (5) & (6)). Where as  $L_{\text{int edge}}$  ensures compatibility of the spatial gradient norm between edge and interior masks themselves.

### 3.4.7. Total Loss

To keep integrity of eight losses used in this model, the total loss estimates normalization coefficient  $N_{\text{loss name}}$  by averaging the values of segmentation, frame field and coupling losses on some part of training dataset. The normalization coefficient helps to rescale the losses and enables them equally contribute to the general loss.

$$\begin{aligned} & \lambda \left( \frac{L_{\text{int}}}{N_{\text{int}}} + \frac{L_{\text{edge}}}{N_{\text{edge}}} + \frac{L_{\text{align}}}{N_{\text{align}}} \right) \\ & + (1 - \lambda) \left( \frac{L_{\text{align } 90}}{N_{\text{align } 90}} + \frac{L_{\text{smooth}}}{N_{\text{smooth}}} + \frac{L_{\text{int align}}}{N_{\text{int align}}} + \frac{L_{\text{edge align}}}{N_{\text{edge align}}} + \frac{L_{\text{int edge}}}{N_{\text{int edge}}} \right) \end{aligned} \quad (11)$$

$\lambda \in [0,1]$  is the only parameter needed to combine the total loss. The value of  $\lambda$  is set to be 0.75 in order to increase the weight of main loss instead of regularizers.

### 3.5. Polygonization Method

Providing vector-based data representations is a challenging task but very critical requirement for multiple geospatial applications (Zorzi et al., 2021). The stand-alone deep learning model by itself does not deliver the segmentation result in a vector format. Thus, additional algorithms are needed to extract object contour from the segmentation map and at a later stage to create regularized outlines.

#### 3.5.1. Active Contour Model (ACM)

The polygonization method is based on the concept of active contour model (ACM) also known as snakes. A snake is a computer-generated curve that recursively moves within the image searching for object boundaries (Kass et al., 1988). Given an initial contour, ACM is an energy-minimizing spline that is guided by external constraints and pushed towards nearby edges and lines. Image data such as gradient magnitude defines an external force that enables initial curve to easily adjust itself to the local minima, while fitted curves define an internal force that determines the elasticity and smoothness of the contour (Bong et al., 2016). In our case, snakes iteratively deform their shape to align to a feature. These features are general boundaries that enclose a parcel of land.

#### 3.5.2. Polygon simplification

The post-processing step, called polygon simplification, reduces the number of vertices in the predicted polygon. It returns a simplified version of the predicted geometry based on the specified tolerance. The tolerance value is a distance. Roughly, any squiggle in a curve that differs from a straight line by less than the tolerance value will be straightened out. The algorithm finds the most extreme deviations that are beyond the tolerance value, pin down those vertices where they deviate, and then recursively rebuild the optimal (straight) line between the pinned-down vertices. The ACM takes the interior classification map generated from the first network head as input and iteratively optimizes a contour to align perfectly on the frame field. This alignment is followed by finding corners and removing vertices that deviate from two ends of corners.

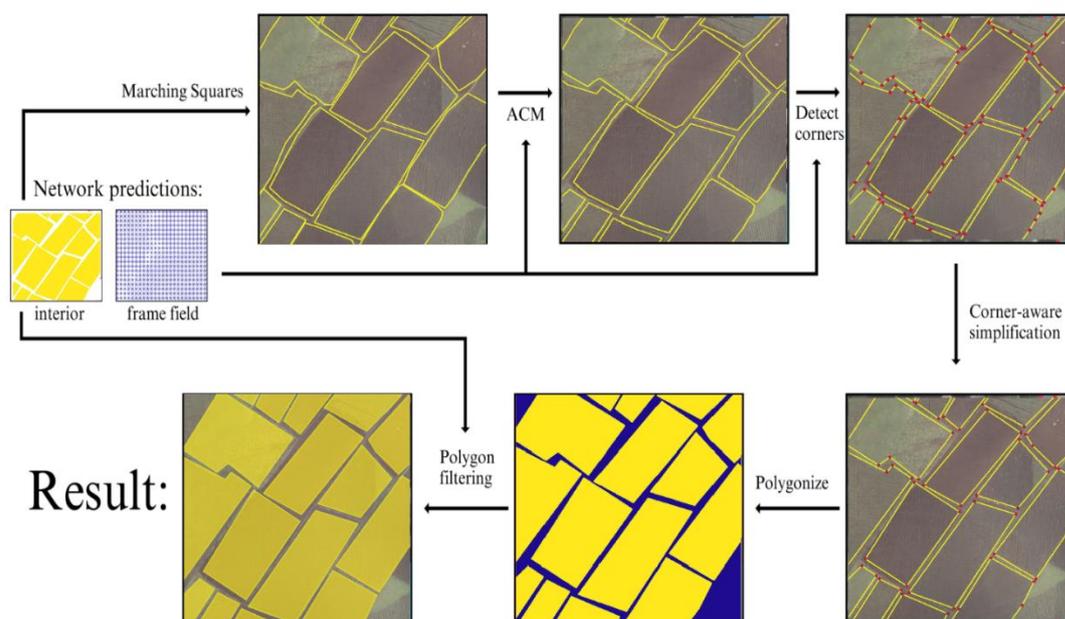


Figure 12: Post-processing and polygonization workflow [adapted from Girard et al. (2020)]

### 3.6. Model Evaluation

The choice of suitable accuracy metrics is crucial when evaluating the model's performance. In this research, pixel and polygon level evaluation metrics are used to evaluate performance of the proposed model results.

#### 3.6.1. Intersection over Union (IoU)

IoU is the preferred evaluation metric for models trained with strong pixel-level semantic segmentation. IoU quantifies the percentage of overlap between the target mask and the predicted polygon (Zhang et al., 2021). This metric is closely related to the Dice coefficient which is often used as a loss function during training. Given a prediction mask  $A$  and the reference data mask  $B$ , the IoU is simply calculated as:

$$IoU(A, B) = \frac{A \cap B}{A \cup B} \quad (12)$$

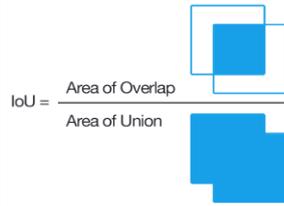


Figure 13: Pixel-wise evaluation using intersection over union metric

#### 3.6.2. Polygons and Line Segments (PoLiS)

Most quantitative evaluation metrics such as the IoU are based upon pixel-by-pixel similarity measures. As a result, these algorithms cannot evaluate the geometric properties of a vector object. Polygons and line segments similarity metric (PoLiS) originally introduced by Avbelj et al.(2015) overcome this limitation by accounting changes in rotation, translation, and scale of the predicted vector object with respect to the reference polygon. Polygons are considered as a sequence of connected vertices instead of standalone points. Given a predicted polygon  $A$  and the reference polygon  $B$ , the PoLiS distance is calculated:

$$\vec{p}(A, B) = \frac{1}{2q} \sum_{aj \in A} \min_{b \in \partial B} \|aj - b\| + \frac{1}{2r} \sum_{bk \in B} \min_{a \in \partial A} \|bk - a\| \quad (13)$$

Where,  $\vec{p}(A, B)$  is the mean distance between all vertices of the predicted polygon and its nearest point (vertex) on the reference polygon. The smaller the PoLiS distance indicates better similarity between predicted and reference polygon.

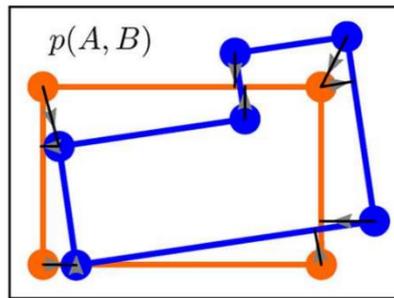


Figure 14: PoLiS distance indicated in black line between extracted (orange) and reference polygon vertices (blue). (Avbelj et al., 2015).

### 3.7. Experimental Setup

The deep learning model is implemented in PyTorch framework using python programming language. The model is trained with Adam optimizer and a batch size of 4, gamma value of 0.95, and learning rate of  $1e-5$ . Moreover, additional experiments is conducted to test transferability of the model for 200 epochs with all setting unchanged except we removed the frame field. It takes 5 minutes per epoch to perform one complete pass. All experiments were conducted on NVIDIA TITAN X (Pascal) 12 GB CUDA GPU remote Linux computer.

### 3.8. Summary

An explanation of the methods to achieve the research objectives is provided in this chapter. A study area is selected in rural part of Ethiopia based on the availability of VHR remote sensing images and cadastral data. The UAV images are processed using WebODM command line utility to create an orthophoto mosaic. The orthophoto is further divided into smaller patches to provide appropriate size into the computer's memory. UNET-ResNet101 network architecture is used as a backbone to perform the segmentation task. Simultaneously the frame filed learning method improves cadastral boundary extraction by adding structural information to a standard segmentation model. Guided by this information, The ACM method iteratively optimizes a contour on predicted segmentation mask with simplification tolerances to deliver vector polygons with regular edges and sharp corners. Finally, two evaluation metrics are introduced to evaluate the model result.

## 4. RESULTS AND DISCUSSION

In this chapter, the cadastral boundaries obtained using the frame field model using the two types of image sources UAV and aerial orthophotos are compared. The model is trained for 200 epochs on both UAV and aerial orthophotos. To make comparisons easier, both scenarios use the same model parameters. To determine the optimum settings for the model, repeated experiments has been conducted. The hyperparameter selection was determined based on lowest validation loss obtained during multiple trials. The final implementation of training hyper-parameters are listed below in Table 2:

Optimizer	BCE coefficient	Dice coefficient	Learning rate	Data augmentation	Batch size	Tolerance	Max epochs
Adam	0.75	0.25	1e-5	Used during training	4	4 pixels	200

Table 2: Training parameters of frame field learning model

### 4.1. Quantitative analysis

All experimental results are presented in the following section. The prediction results are evaluated in pixel wise metric reported in Mean IoU (mIoU) and polygon level metric reported in PoLiS distance value. The first experiment was conducted on 61 test tiles. The values obtained on each tile is averaged and presented in Table 3.

Orthophoto type	Average PoLiS distance	mIoU
UAV	2.81	0.84
Aerial	8.64	0.79

Table 3: Summary of frame field learning model prediction results

Due to the higher spatial resolution, a higher Mean IoU of 0.84 was achieved on UAV orthophoto compared to aerial orthophoto (0.79). Even though the model benefits from high resolution UAV orthophoto, it was also susceptible to oversegmentation. Tile numbers 4 and 5 in Table 4 indicate that the total number of matches (predicted polygons) exceeds the total number of polygons in the reference data.

It's also important to examine the similarity between the predicted and reference polygons since this indicates the quality of model results. Polygons predicted using the UAV orthophoto dataset have an average PoLiS distance of 2.81, four times lower than those predicted using the aerial orthophoto dataset, where the average PoLiS distance is 8.64. The lower the PoLiS distance is the better. This result indicates UAV-based polygon predictions are more similar with the reference data than aerial-based polygon predictions.

For simplicity and visualization, we picked out seven tiles from both UAV and aerial test datasets to show detailed cases on each tile.

Tile No.	Orthophoto type	Total no. of parcels in reference data	Number of matches	Number of misses	Duplicate matches	PoLiS Distance
1	UAV	14	14	0	2	1.14
	Aerial		11	3	1	4.19
2	UAV	16	16	0	1	2.10
	Aerial		16	0	1	3.96
3	UAV	16	19	-3	4	2.49
	Aerial		1	15	0	29.21
4	UAV	15	18	-3	4	5.46
	Aerial		13	2	2	7.32
5	UAV	10	10	0	0	2.43
	Aerial		10	0	1	4.50
6	UAV	14	13	1	1	3.42
	Aerial		5	9	0	5.90
7	UAV	8	9	-1	2	2.95
	Aerial		8	0	2	5.43

Table 4: Quantitative results obtained on UAV and aerial test datasets

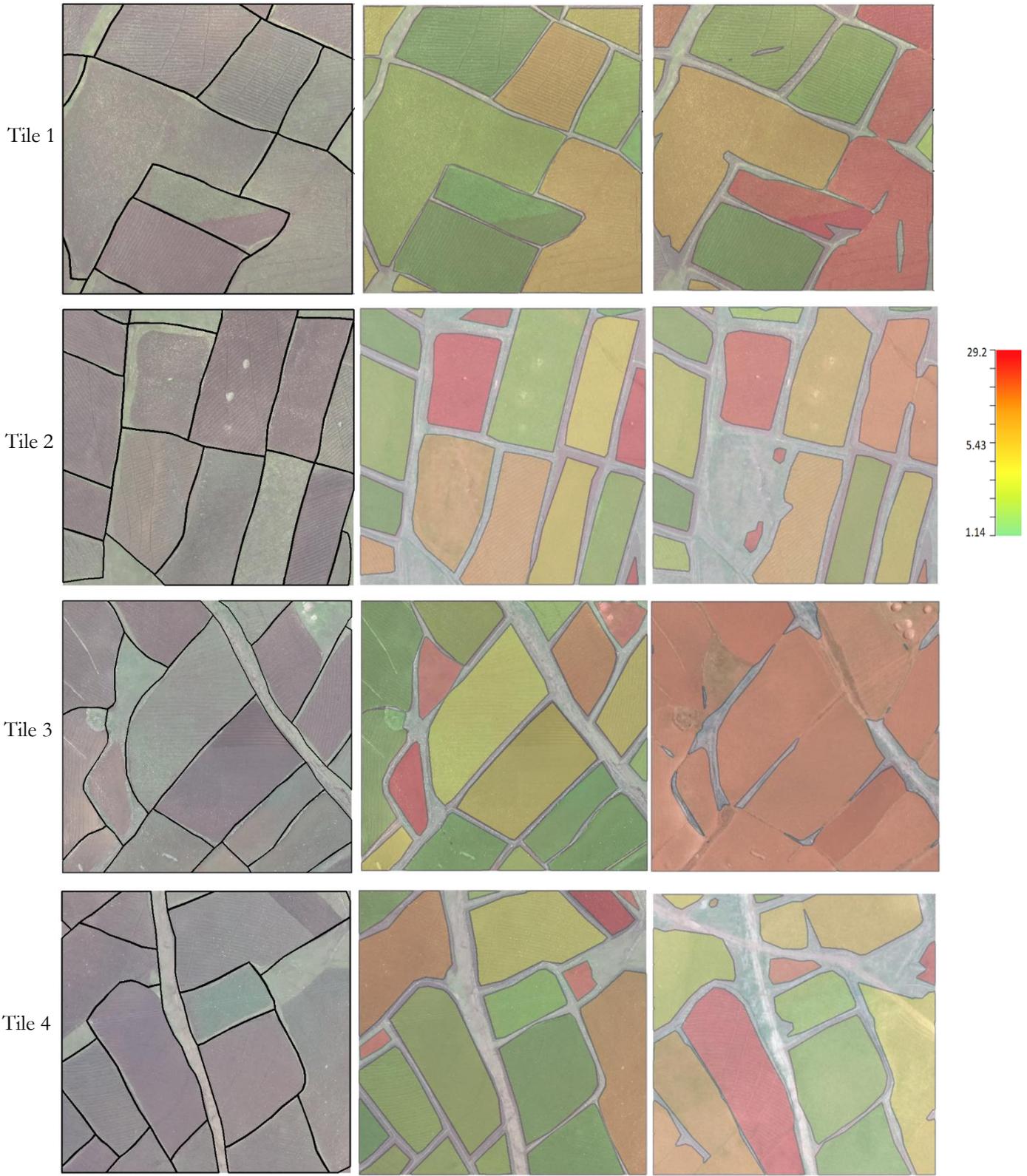
In the above table, the number of matches correspond to all candidate polygons predicted by the model whereas the number of misses correspond to not predicted polygons and false positive results. Duplicate matches indicate the number of times a single polygon was predicted more than once.

There are more matches in the UAV-based predictions than in the aerial orthophoto dataset, on the contrary predictions made on aerial orthophoto dataset have the largest number of missed polygons. More duplicated polygons are found in the predictions of the UAV orthophoto dataset than in those of the aerial dataset. This is related to the over segmentation problem that occurred due to the high-resolution nature of UAV orthophoto. Negative numbers on the the above table represent extra polygons predicted on UAV dataset.

#### 4.2. Qualitative analysis

In Figure 15, the predicted polygons are visualized with a graduated color scale based on to their PoLiS distance (similarity) value. The dark green polygons represent the best polygon prediction examples whereas the red end indicates the worst prediction results. The PoLiS distance value reveals the correlation between the reference and predicted polygons. In other words, the smaller PoLiS distance value, the higher the similarity between the reference and predicted polygons and vice versa. Polygons predicted on UAV orthophoto are more aligned with the reference data than polygons predicted on aerial orthophoto. qualitative results of selected tiles are shown below in

Figure 15.



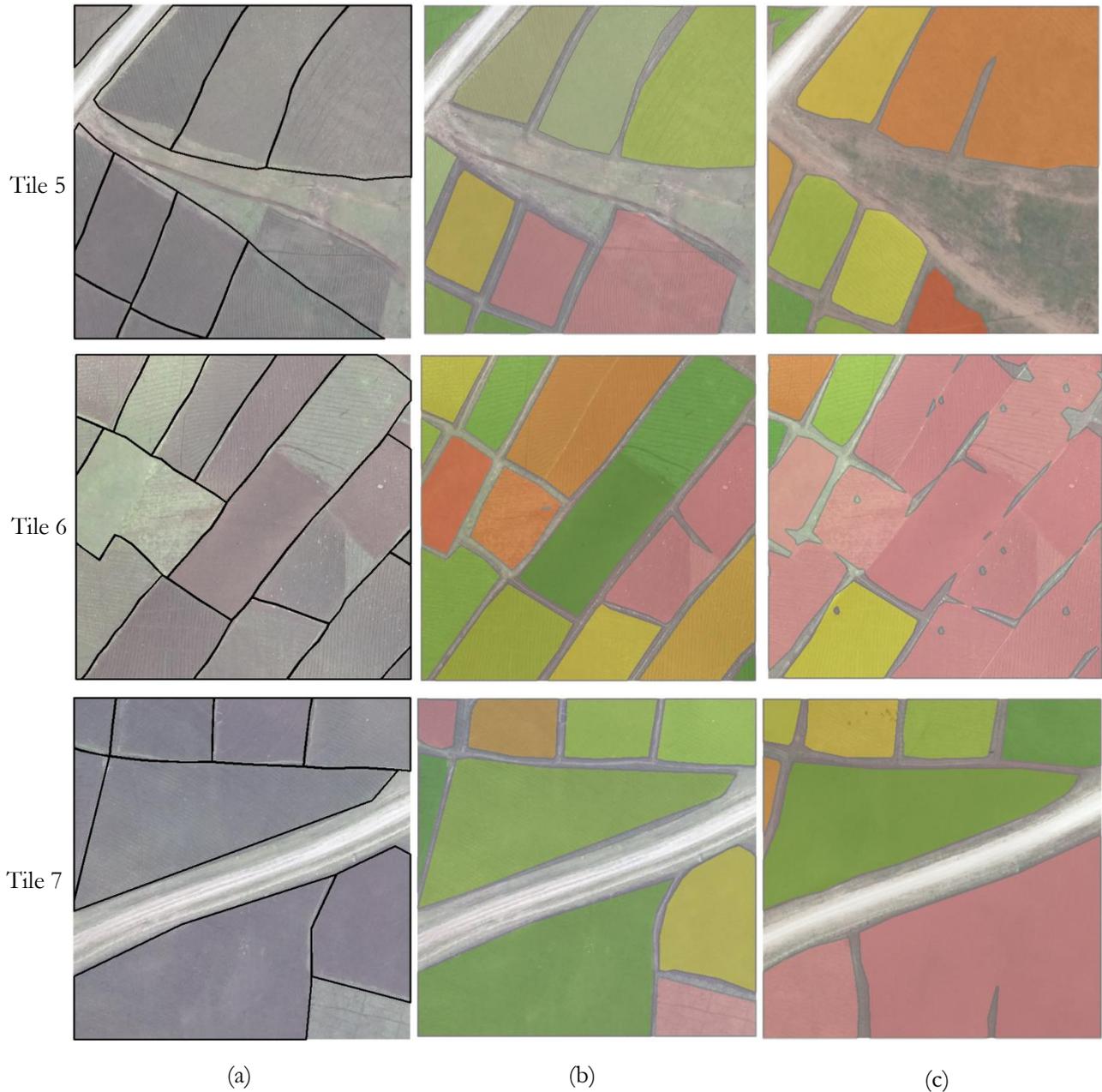


Figure 15: reference data overlaid on the UAV orthophoto (a) predicted polygons on the UAV orthophoto (b) predicted polygons on aerial orthophoto (c)

Generally, polygons predicted with UAV orthophoto have shown higher similarity with the reference data than those predicted with the aerial orthophoto. Polygon predictions based on aerial orthophoto have shown a major drawback of under segmentation. Many of the predicted polygons are merged together due to similar spectral information. Another aspect that generally contributed to higher PoLiS distance value is quality of the reference data. In some tiles, the reference polygon size gets bigger or smaller than the actual parcel size due to errors committed when digitizing the boundaries by human operators.

### 4.3. Over and Under-Segmentation

One of the most common problems in segmentation is that there are too few (under segmented) or too many (over segmented) regions. In our experiment, we found a few cases where the resolution of images played a role in over and under segmentation. As the UAV orthophoto have higher spatial resolution, it contains a range of textures, leading to over segmentation. Whereas under segmentation was observed on polygons predicted from the aerial orthophoto. Besides the resolution difference, the aerial orthophoto was acquired in a season when most of the vegetation was dry, and no agricultural activity took place. Such condition creates fuzzy spectral characteristics between fields making detection of boundary difficult.

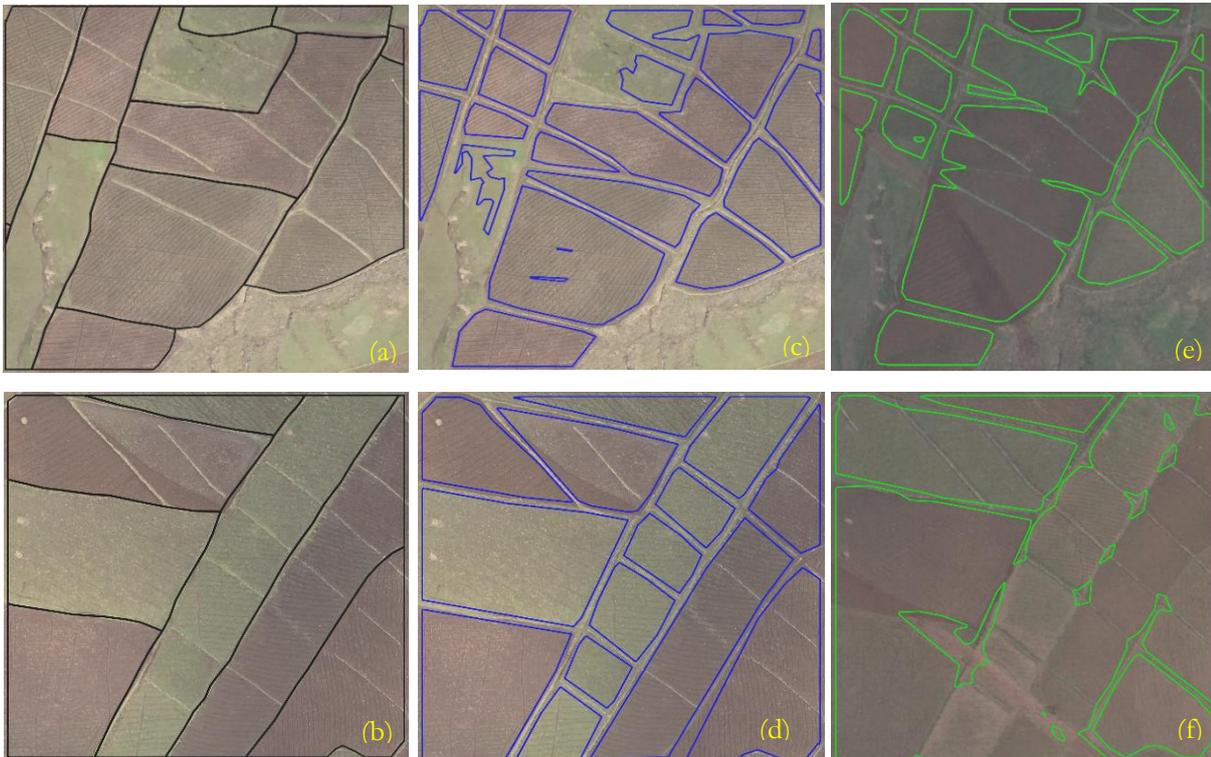


Figure 16: Over and under-segmentation cases: reference data overlaid on UAV orthophoto (black) prediction on UAV orthophoto (blue) prediction on aerial orthophoto (green)

Cases of over and under segmentation are shown in Figure 16. The original reference data in (a) contains a total of 11 reference polygons. The model prediction on UAV orthophoto (c) returned twice the original number of polygons in the reference data. This is due to the fact that crops can vary within one plot of land which causes the model to generate false positive results. Small waterways that run through agricultural fields also cause confusion to the model. On the other hand, The original reference data in (b) contains a total of 7 reference polygons and the model prediction on aerial orthophoto (f) returned 4 polygons that clearly indicates undersgmentation.

For manual post-processing, it is easier to join small segments to form a larger polygon rather than splitting a large segment into smaller polygons.

Some instances of under segmentation are related to the existence of invisible cadastral boundaries. As shown in Figure 17, the reference data (b) contains two separate holding rights (indicated by white arrows). Because the boundary that separates these two holding rights is invisible (perceived), the model returned a merged polygon (c) treated as one holding right.

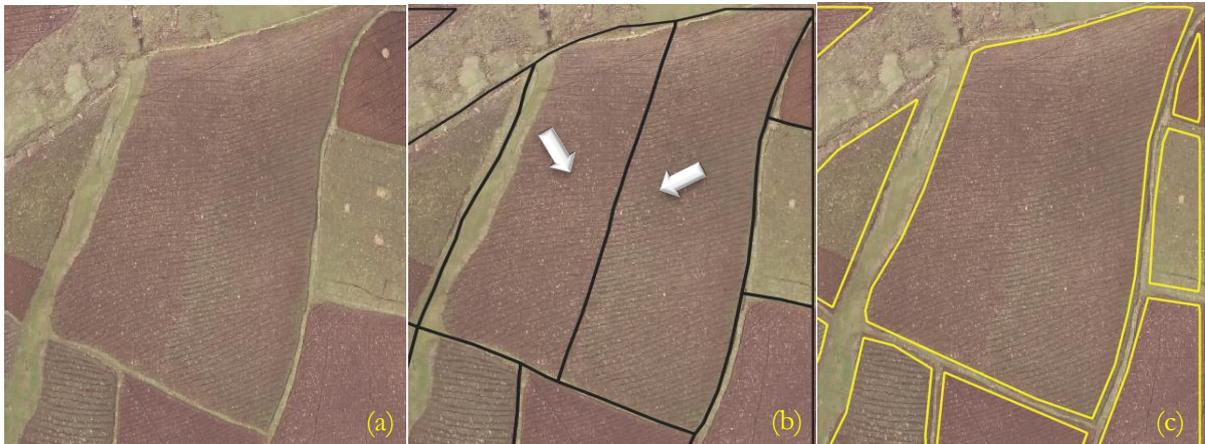


Figure 17: Prediction on invisible boundaries: raw image (a) perceived holding rights indicated by the white arrows (b) predicted boundary (c)

Considering that the reference cadastral boundaries were digitized on top of the aerial images that was acquired in 2017, there are two years of difference between the aerial and UAV orthophotos. It is clearly visible in the UAV orthophoto (Figure 18c) that there have been recent updates on the ground.



Figure 18: Detected boundary changes on the UAV orthophoto

In Figure 18, the triangle-shaped parcel indicated by blue arrow (b) is the original reference data overlaid on top of the aerial image. A footpath borders the right side of the property indicated by red arrow (a). Over time, however, the morphology (land use) has changed, including the footpath indicated by the white arrow and boundary of the parcel indicated by the yellow arrow (c). Detecting updated boundaries was made possible by the model after being able to observe these changes clearly in the UAV image.

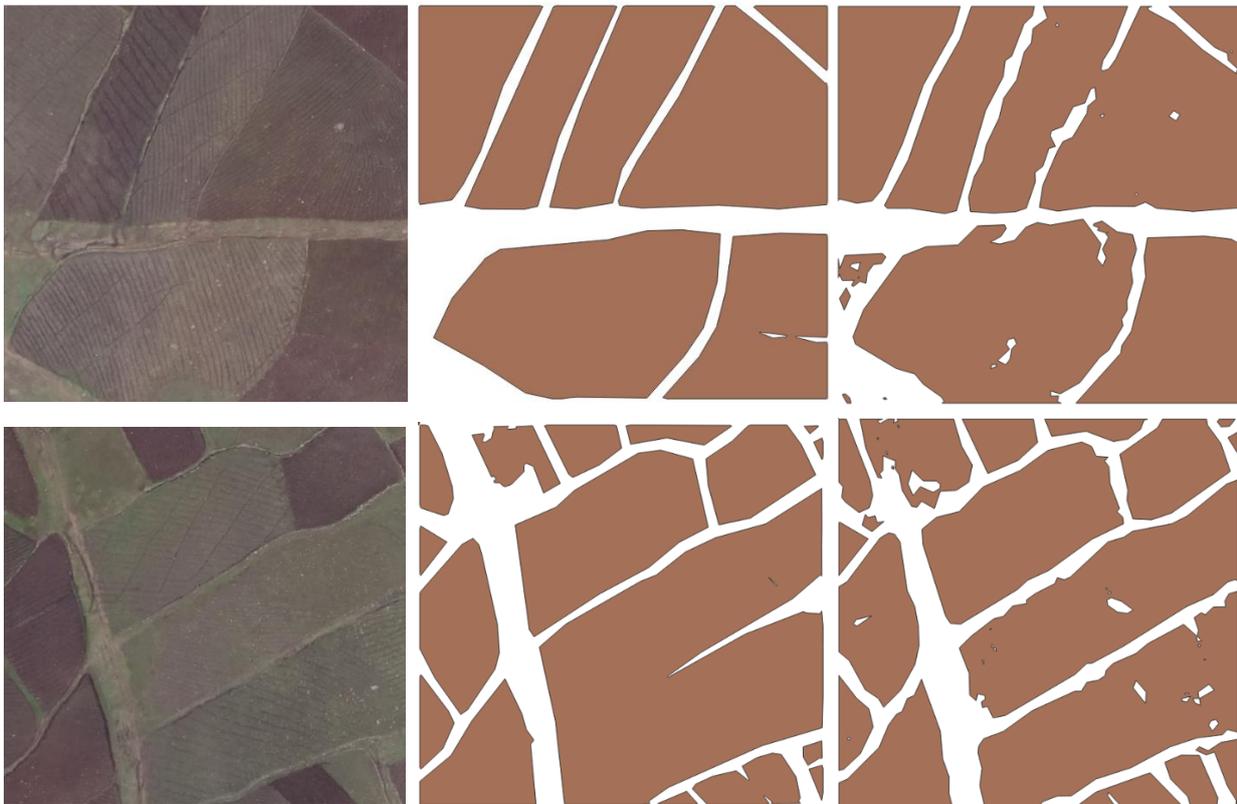
#### 4.4. Comparison of frame field results with standard segmentation model

Using the UAV orthophoto, additional model training was performed by removing the frame field and only using the segmentation masks to explore the effect of absence of frame field losses and regularizers on polygon edge and corners. Four tiles were prepared to demonstrate the difference. The IoU and PoLiS distance values are calculated by taking the average.

Backbone Network	Polygonization Method	Tolerance	mIoU	Average PoLiS Distance
UNET-ResNet101 (no field)	Simple	4 pixel	0.81	4.75
UNET-ResNet101 (with frame field)	ACM	4 pixel	0.83	4.54

Table 5: Comparative test result of frame field and standard segmentation model

As presented in Table 5, there is a slight difference in IoU and PoLiS distance values of the two models. But qualitatively it is clearly visible that the model trained with frame field have benefited from regular regular edges and corners.



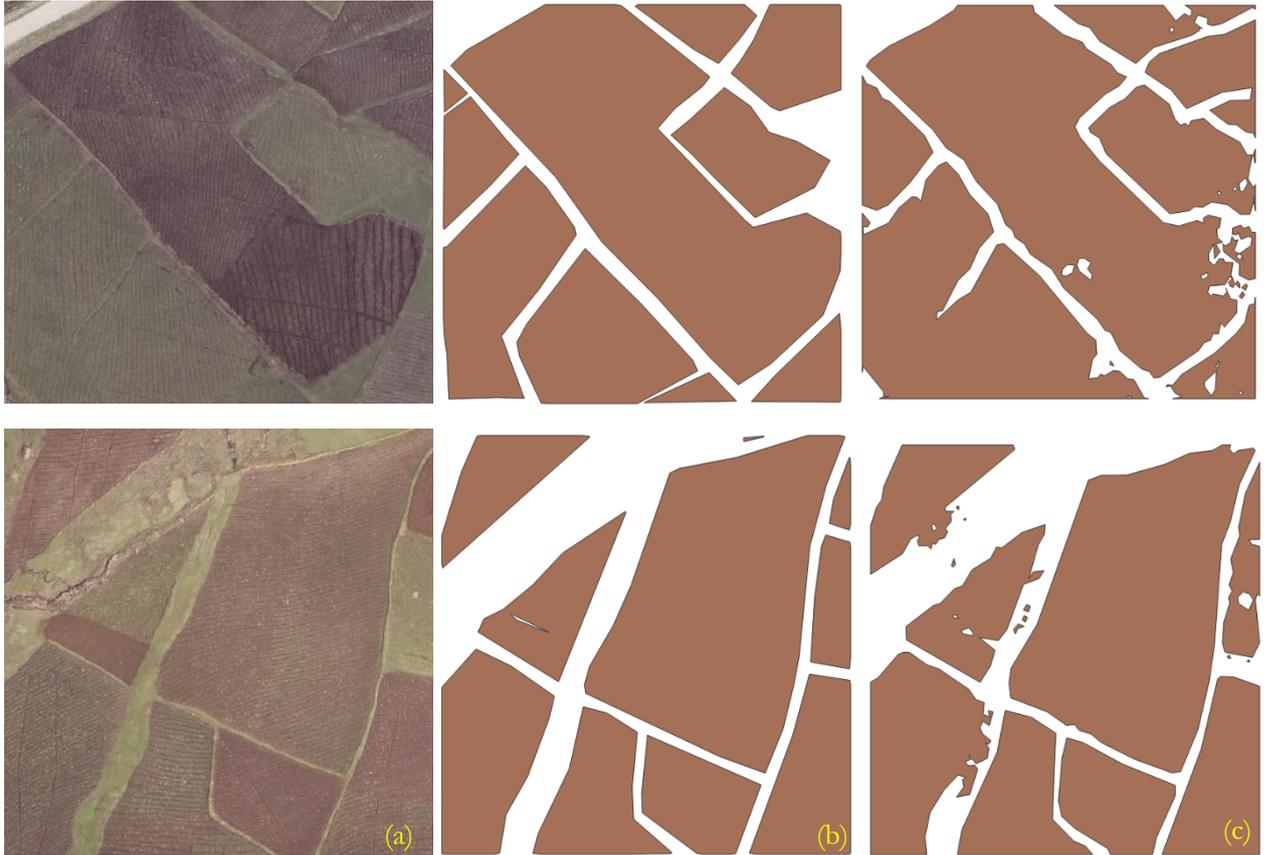
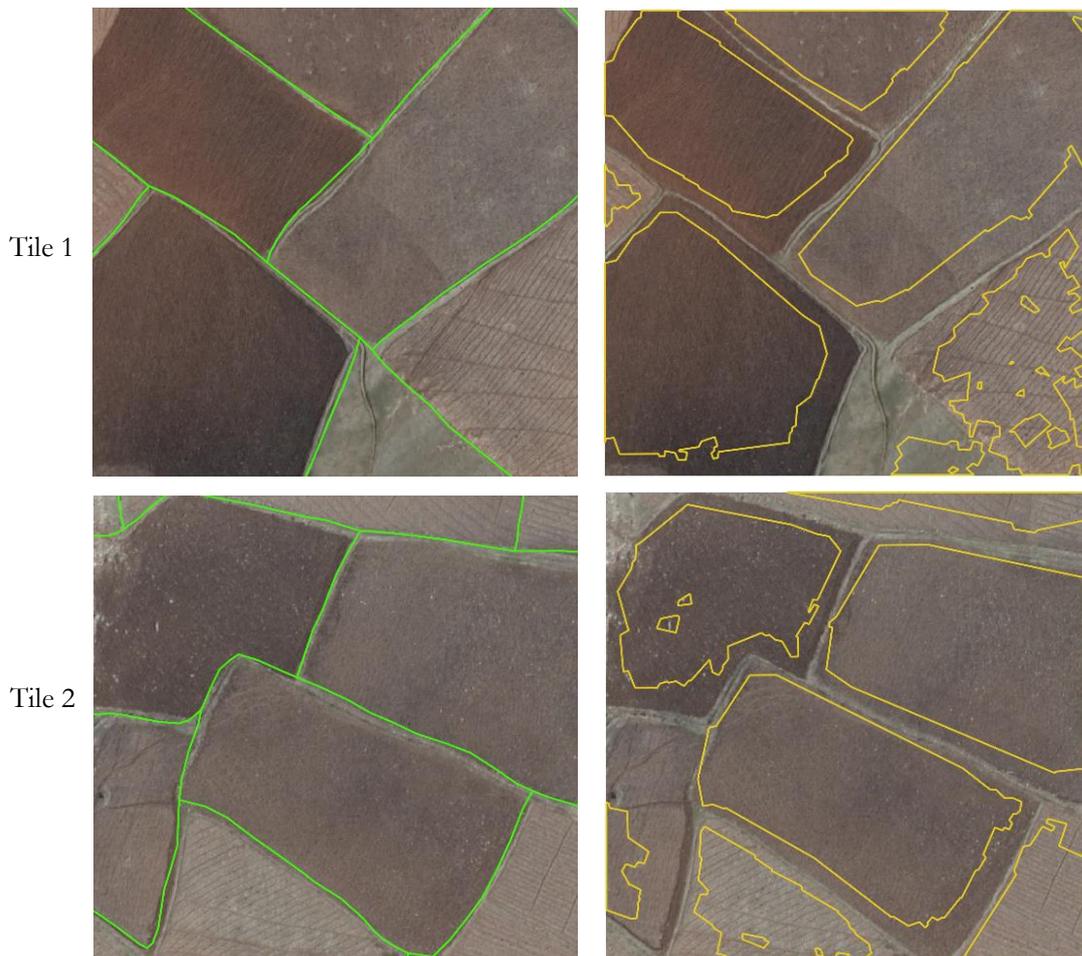


Figure 19: Comparative results: Raw image (a), U-NetResNet101: trained with frame field (b) U-NetResNet101: trained without frame field (c)

In Figure 19, prediction results based on U-NetResNet101 trained without frame fields using simple polygonization method are characterized by irregular shaped edges. The absence of frame field losses and regularizers has a clear impact on the polygonization stage. The 8 losses in the frame field learning model, of which one of them serve as regularizer, plays a significant role by delivering simplified and clean polygons.

#### 4.5. Testing transferability of the model

To determine if the model can reproduce the same result in a different geographical area, we evaluated the model by preparing additional test dataset using the aerial orthophoto, which is separate from the previous dataset that were used for comparing with UAV results. The experiment was conducted on three tiles and the mIoU achieved is 0.67 which is lower than the previous experimental result achieved on aerial orthophoto (0.79). on the other hand, the average PoLiS distance of 7.11 is obtained which is better than the previous result (8.64) achieved on the aerial orthophoto.



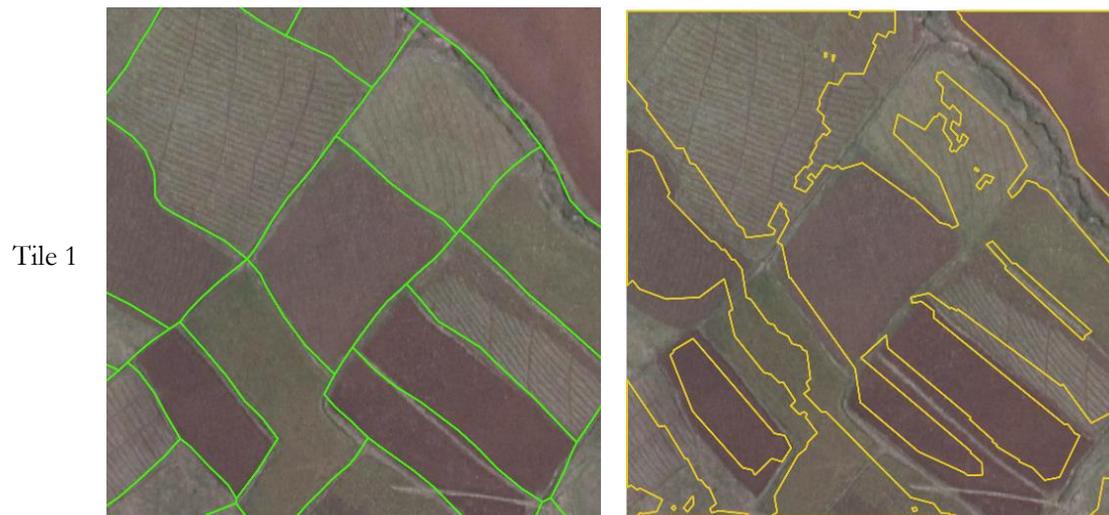


Figure 20: Results obtained using aerial image dataset on different geographical area. Reference data overlaid on aerial orthophoto (green) and model predictions on aerial orthophoto (yellow)

A general observation was that the time of image acquisition plays an important role when planning to use machine learning algorithms for automatic boundary extraction. Due to the fact that the aerial image was taken during the dry season, many agricultural fields have similar and fuzzy spectral characteristics which causes under segmentation.

## 5. CONCLUSION

In this study, a deep learning model combined with frame field learning has been explored and evaluated to extract cadastral boundaries of agricultural fields in rural part of Ethiopia using very high-resolution remote sensing images in vector polygon format. The proposed method utilizes FCN based network specifically UNET as a backbone to perform the segmentation task on the input image followed by the frame field method which simultaneously learns directional information of each pixel on the image. Through multi-task learning, frame field output improves the quality of deep segmentation model while providing structural information to facilitate the polygonization process. The ACM polygonization method takes the segmentation mask and frame field as input and iteratively optimizes edges and corners to align seamlessly with the reference data.

The proposed method was evaluated on two different orthophotos based on images taken from UAV and aircraft platforms with 0.11 and 0.29 cm. resolution, respectively. In addition to the standard semantic segmentation metric (IoU), PoLiS metric was introduced which allows us to quantify changes in rotation, translation, and scale of predicted polygon with respect to the reference data. The smaller PoLiS distance value indicates better similarity of predicted polygon with its corresponding reference polygon. According to the quantitative results, polygons predicted on UAV orthophotos have higher similarity with the reference polygon than polygons predicted on aerial orthophotos with PoLiS distance of 2.81 and 8.64, respectively. Furthermore, a higher mean IoU of 0.84 was achieved on polygons predicted using UAV orthophotos compared to predictions on aerial orthophotos which is 0.79.

A standard segmentation model was compared with our frame-field based model to assess the effects of regularization losses on predicted polygons. The results revealed that frame field based model predictions delivered simplified polygons with regular edge and corner. Additional test was done in another geographical area using aerial orthophoto to check transferability of our model. The model achieved mIoU of 0.67 and PoLiS distance of 7.11 which is slightly lower than the quality of polygons predicted using UAV orthophoto.

The quality of the predicted boundaries has been compromised by a number of factors including high crop variability inside one field, existence of invisible boundaries, confusing features such as water ways and terraces that crosses the parcel, date of image acquisition and the quality of reference cadastral data.

In conclusion, the proposed method shows the possibility of utilizing deep learning methods for extracting cadastral boundaries in a vector polygon format that can be directly used in mapping applications with little post-processing.

## 5.1. Reflection to research objectives and questions

The main purpose of this this thesis is to design a deep learning method that can extract cadastral boundaries in a vector polygon format using VHR remote sensing images which aims to reduce the time and cost of data collection on the ground. In the following section we provide summarized answer to the sub objectives and questions that we have asked in chapter one.

### 5.1.1. Objective 1: To investigate deep learning methods used for cadastral boundary extraction

- i. What methods exist for cadastral boundary extraction

A systematic literature review has been conducted on previous studies that has implemented deep learning algorithms to extract cadastral boundaries using VHR remote sensing images. Fully convolutional networks (FCN) have found to be highly effective for semantic segmentation. This method outperforms other state of art methods such as global probability of boundary (gPb), ultrametric contour map (UCM), and oriented watershed transform (OWT). In total, three variants of FCN has been reviewed and tested, including FCNDK, UNET, and SegNet.

### 5.1.2. Objective 2: To prepare data for cadastral boundary extraction

- i. What are the steps needed to post-process UAV images?

Raw images captured by the UAV vehicle were orthorectified to produce geometrically corrected orthophoto. This was achieved by using open drone map (ODM) software, an open source command line utility designed to process photogrammetric products free of cost. Using the GCPs and intrinsic and extrinsic parameters of the camera, ODM automatically creates the orthophoto.

- ii. How to design & prepare input dataset for the model?

Training, validation, and test datasets are prepared in the area where UAV and aerial orthophotos overlapped. The reason for choosing overlapped area is to ensure fair comparison of results from two different orthophotos acquired at the same location. To overcome computing resource issues, training and validation datasets are further sub divided into 512x512 pixels using image tiling technique. To evaluate the model on larger area, the test dataset patch size is set to 1024x1024 pixels. For each image tile, the corresponding reference data is prepared from the official cadastral boundaries of the study area. This task was completed using Solaris, an open source python library used for geospatial machine learning projects.

### 5.1.3. Objective 3: To design a deep learning method used for boundary extraction

- i. Which FCN based network architecture is suitable to use as backbone?

FCNDK, deeplapv3 SegNet, and UNET were compared to determine which FCN-based network would be the backbone of our method, and UNET delivered the best segmentation results. Additionally, the segmentation results have greatly improved when the encoder path of UNET is replaced with a pretrained ResNet101 encoder.

ii. How to get optimum hyperparameters of the model?

Several hyperparameters such as the optimizer, learning rate, batch size, loss function, and polygon simplification tolerances are compared by performing experimental runs. The optimal model hyperparameters are then selected based on the epoch with the lowest validation loss.

iii. What methods are used to extract regularized vector polygons?

The frame filed learning method improves cadastral boundary extraction by adding structural information to a standard segmentation model. Guided by this information, The ACM method iteratively optimizes a contour on predicted segmentation mask with simplification tolerances to deliver vector polygons with regular edges and sharp corners.

#### 5.1.4. Objective 4: To evaluate the performance of the method

i. What are the standard metrics used for evaluating predicted polygons?

The standard metric used for evaluating semantic segmentation is Intersection over Union (IoU). But IoU does not consider the geometric properties of predicted polygons. Therefore, we have introduced PoLiS metric which allows us to quantify changes in rotation, translation, and scale of predicted polygon with respect to the reference data. Smaller PoLiS distance gives indication of optimal prediction result.

ii. What is the performance of the model on UAV and aerial orthophoto?

The prediction results are compared on UAV and aerial orthophoto. Polygons predicted on UAV orthophoto have higher similarity with their corresponding reference polygon when compared to polygons predicted on aerial orthophoto. Based on to the quantitative results, polygons predicted on UAV orthophotos have higher similarity with the reference polygon than polygons predicted on aerial orthophotos with PoLiS distance of 2.81 and 8.64, respectively. Furthermore, a higher mean IoU of 0.84 was achieved on polygons predicted using UAV orthophotos compared to predictions on aerial orthophotos which is 0.79. In addition, the model is tested in another geographical area using aerial orthophoto to test its transferability and mIoU of 0.67 and PoLiS distance of 7.11 is achieved which is slightly lower than the quality of polygons predicted using UAV orthophoto.

## 6. RECOMMENDATION

Based on current research, recommendations for future works are listed as follows:

1. Future works should investigate the use of Active skeleton Model (ASM) to obtain polygons that share the same boundary. This polygonization method is introduced by Girard et al. (2020) and tested on buildings with shared walls in urban areas. The transferability of this polygonization method could be tested in rural areas (agricultural fields).
2. The use of multispectral data can increase the quality of image segmentation. Future research should consider fusing additional bands with RGB orthophoto.
3. A planned image acquisition should consider the optimal time of the year for capturing rich spectral information.



## LIST OF REFERENCES

---

- Aung, H.L., Uz Kent, B., Burke, M., Lobell, D., Ermon, S., 2020. Farm Parcel Delineation Using Spatio-temporal Convolutional Networks, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, pp. 340–349. <https://doi.org/10.1109/CVPRW50498.2020.00046>
- Avbelj, J., Muller, R., Bamler, R., 2015. A Metric for Polygon Comparison and Building Extraction Evaluation. *IEEE Geoscience and Remote Sensing Letters* 12, 170–174. <https://doi.org/10.1109/LGRS.2014.2330695>
- Ayano, M.F., 2018. Rural Land Registration in Ethiopia: Myths and Realities. *Law & Society Review* 52, 1060–1097. <https://doi.org/10.1111/lasr.12369>
- Badrinarayanan, V., Handa, A., Cipolla, R., 2015. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling.
- Blachut, T.J., 1976. The Stereo-Orthophoto Technique in Cadastral and General Mapping. *Photogramm. Eng. Remote Sensing* 42, 1511–1519.
- Bessmeltsev, M., Solomon, J., 2018. Vectorization of Line Drawings via PolyVector Fields. *ACM Transactions on Graphics* 38, 1–12. <https://doi.org/10.1145/3202661>
- Bezu, S., Holden, S., 2014. Demand for second-stage land certification in Ethiopia: Evidence from household panel data. *Land Use Policy* 41, 193–205. <https://doi.org/10.1016/j.landusepol.2014.05.013>
- Bong, C.W., Liew, C.C., Lam, H.Y., 2016. Ground-glass opacity nodules detection and segmentation using the snake model. *Bio-Inspired Computation and Applications in Image Processing* 87–104. <https://doi.org/10.1016/B978-0-12-804536-7.00005-3>
- Byamugisha, F.F.K., 2013. Securing Africa’s Land for Shared Prosperity, Africa Development Forum. The World Bank. <https://doi.org/doi:10.1596/978-0-8213-9810-4>
- Cochrane, L., Hadis, S., 2019. Functionality of the Land Certification Program in Ethiopia: Exploratory Evaluation of the Processes of Updating Certificates. *Land (Basel)* 8, 1–14. <https://doi.org/10.3390/land8100149>
- Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 92, 79–97. <https://doi.org/10.1016/j.isprsjprs.2014.02.013>
- Corlazzoli, M., Fernandez, O.L., 2004. SPOT 5 Cadastral validation project in Izabal, Guatemala. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. Part 1*.
- Crommelinck, Koeva, Yang, Vosselman, 2019. Application of Deep Learning for Delineation of Visible Cadastral Boundaries from Remote Sensing Imagery. *Remote Sensing* 11, 1–21. <https://doi.org/10.3390/rs11212505>
- Crommelinck, S., Bennett, R., Gerke, M., Nex, F., Yang, M., Vosselman, G., 2016. Review of Automatic Feature Extraction from High-Resolution Optical Sensor Data for UAV-Based Cadastral Mapping. *Remote Sensing* 8, 689. <https://doi.org/10.3390/rs8080689>
- DAI, 2020. Ethiopia—Land Investment for Transformation (LIFT) · DAI: International Development [WWW Document]. URL <https://www.dai.com/our-work/projects/ethiopia-land-investment-transformation-lift> (accessed 11.22.21).
- De Soto, H., 2000. *The Mystery of Capital : Why Capitalism Triumphs in the West and Fails Everywhere Else*. Basic Books, New York.
- Dale, P.F., 1977. Cadastres and Cadastral Maps. *Cartogr. J.* 14, 44–48. <https://doi.org/10.1179/caj.1977.14.1.44>
- Diamanti, O., Vaxman, A., Panozzo, D., Sorkine-Hornung, O., 2015. Integrable PolyVector fields. *ACM Transactions on Graphics* 34. <https://doi.org/10.1145/2766906>

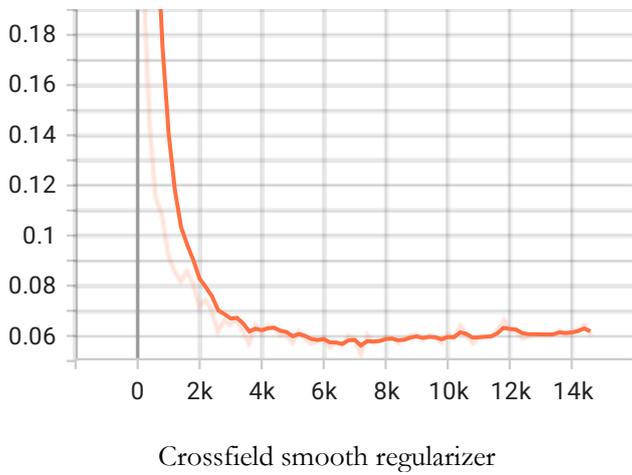
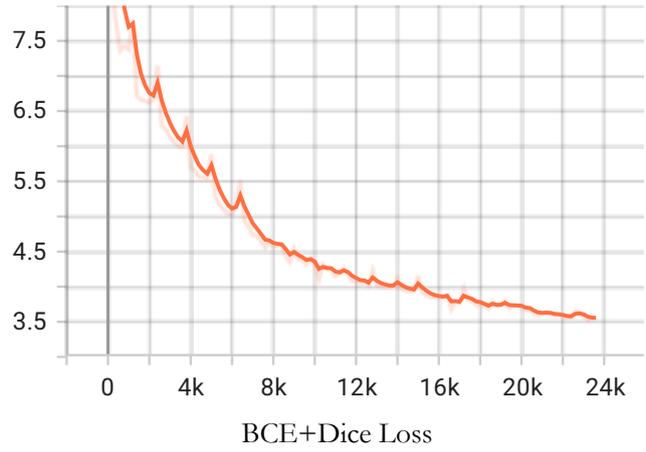
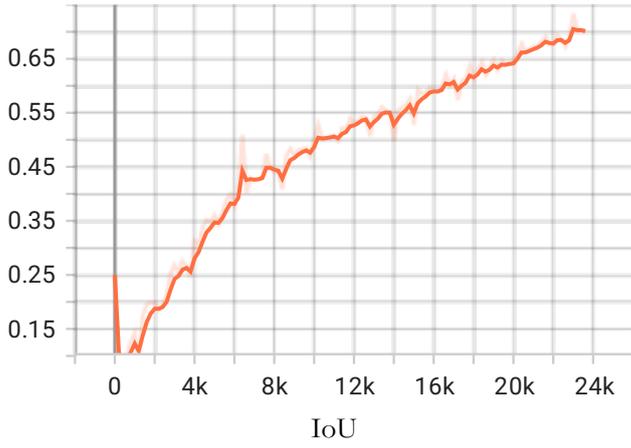
- Diamanti, O., Vaxman, A., Panozzo, D., Sorkine-Hornung, O., 2014. Designing N -PolyVector Fields with Complex Polynomials. *Computer Graphics Forum* 33, 1–11. <https://doi.org/10.1111/cgf.12426>
- Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The Importance of Skip Connections in Biomedical Image Segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 10008 LNCS, 179–187. <https://doi.org/10.48550/arxiv.1608.04117>
- Eisenbeiss, H., 2011. The Potential of Unmanned Aerial Vehicles for Mapping, in: *Photogrammetry Week 2011*. Berlin & Offenbach, pp. 135–145.
- Eker, O., Seker, D.Z., 2008. SEMI-AUTOMATIC EXTRACTION OF FEATURES FROM DIGITAL IMAGERY. *ISPRS J. Photogramm. Remote Sens.* XXXVII, 443–446.
- Enemark, S., Bell, C., Lemmen, C., McLaren, R., 2014. Fit-For-Purpose Land Administration, The International Federation of Surveyors (FIG).
- Enemark, S., McLaren, R., Lemmen, C., Antonio, D., Gitau, J., De Zeeuw, K., Dijkstra, P., Quinlan, V., Freccia, S., 2016. fit-for-purpose land administration guiding principles for country implementation editing and layout: fit-for-purpose land administration guiding principles for country implementation [WWW Document]. URL [https://www.fig.net/news/news\\_2016/2016\\_07\\_gltnguide/fit-for-purpose-land-admin-guiding-principles%02for-country-implementation.pdf](https://www.fig.net/news/news_2016/2016_07_gltnguide/fit-for-purpose-land-admin-guiding-principles%02for-country-implementation.pdf) (accessed 5.26.22).
- Enemark, S., McLaren, R., 2018. Making FFP Land Administration Compelling and Work in Practice, in: *FIG Commission 7 2018 Annual Meeting and International Seminar*. International Federation of Surveyors, Burgen, Norway, pp. 1–14.
- FAO, 2002. Land tenure and rural development, Land Tenure Studies. Rome.
- Fetai, B., Račić, M., Lisec, A., 2021. Deep Learning for Detection of Visible Land Boundaries from UAV Imagery. *Remote Sensing* 13, 1–19. <https://doi.org/10.3390/rs13112077>
- FIG, 1995. FIG Statement on the Cadastre. The International Federation of Surveyors (FIG).
- Garcia-Pedrero, A., Lillo-Saavedra, M., Rodriguez-Esparragon, D., Gonzalo-Martin, C., 2019. Deep Learning for Automatic Outlining Agricultural Parcels: Exploiting the Land Parcel Identification System. *IEEE Access* 7, 1–14. <https://doi.org/10.1109/ACCESS.2019.2950371>
- Girard, N., Smirnov, D., Solomon, J., Tarabalka, Y., 2020a. Regularized Building Segmentation by Frame Field Learning, in: *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, pp. 1805–1808. <https://doi.org/10.1109/IGARSS39084.2020.9324080>
- Girard, N., Smirnov, D., Solomon, J., Tarabalka, Y., 2020b. Polygonal Building Segmentation by Frame Field Learning. *arXiv* 1–30.
- GLTN, 2012. *Handling Land: Innovative Tools for Land Governance and Secure Tenure*. UN-Habitat, Nairobi.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G., Cai, J., Chen, T., 2015. Recent Advances in Convolutional Neural Networks.
- He, H., Zhou, J., Chen, M., Chen, T., Li, D., Cheng, P., 2019. Building Extraction from UAV Images Jointly Using 6D-SLIC and Multiscale Siamese Convolutional Networks. *Remote Sensing* 11. <https://doi.org/10.3390/rs11091040>
- Hendriks, B., Zevenbergen, J., Bennett, R., Antonio, D., 2019. Pro-poor land administration: Towards practical, coordinated, and scalable recording systems for all. *Land Use Policy* 81, 21–38. <https://doi.org/10.1016/j.landusepol.2018.09.033>
- Hensen J.L.G, 1995. Cadastre and its legal aspects, in: *Cadastral Information Systems*. JEC-GI, The Hague, Netherlands.
- Hull, S.A., Kingwill, R., Fokane, T., 2020. An Introduction to Land Administration. <https://doi.org/10.13140/RG.2.2.29539.17442>
- IAAO, 2004. *Standard on Manual Cadastral Maps and Parcel Identifiers*. Chicago.

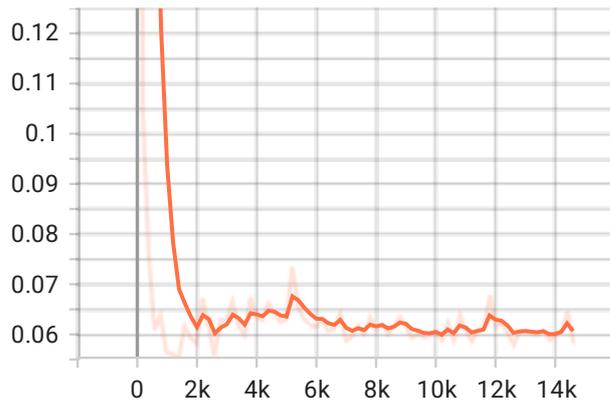
- Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active Contour Models. *International Journal of Computer Vision* 321–331.
- Kelm, K., Antos, S., McLaren, R., 2021. Applying the FFP Approach to Wider Land Management Functions. *Land (Basel)* 10, 1–25. <https://doi.org/10.3390/land10070723>
- Koeva, M., Humayun, M.I., Timm, C., Stöcker, C., Crommelinck, S., Chipofya, M., Bennett, R., Zevenbergen, J., 2021. Geospatial Tool and Geocloud Platform Innovations: A Fit-for-Purpose Land Administration Assessment. *L. 2021, Vol. 10, Page 557* 10, 557. <https://doi.org/10.3390/LAND10060557>
- Kohli, D., Bennett, R., Lemmen, C., Asiama, K., Morales, A., Pinheiro, A., Wayumba, R.N., Zevenbergen, J.A., 2017. A Quantitative Comparison of Completely Visible Cadastral Parcels Using Satellite Images: A Step towards Automation, in: *Surveying the World of Tomorrow - From Digitalisation to Augmented Realty*. International Federation of Surveyors (FIG), pp. 1–14.
- Luo, X., Bennett, R., Koeva, M., Lemmen, C., Quadros, N., 2017. Quantifying the Overlap between Cadastral and Visual Boundaries: A Case Study from Vanuatu. *Urban Science* 1, 32. <https://doi.org/10.3390/URBANSCI1040032>
- Masoud, K.M., Persello, C., Tolpekin, V.A., 2019. Delineation of Agricultural Field Boundaries from Sentinel-2 Images Using a Novel Super-Resolution Contour Detector Based on Fully Convolutional Networks. *Remote Sensing* 12, 1–16. <https://doi.org/10.3390/rs12010059>
- Milindi Rugema, D., Amsalu Birhanu, T., Belay Shibeshi, G., 2020. Evaluating Spatial Data Acquisition and Management Techniques for Multipurpose Cadastre in Ethiopia and Rwanda (Preliminary Results). *African J. L. Policy Geospatial Sci.* 3, 195–212. <https://doi.org/10.48346/IMIST.PRSM/AJLP-GS.V3I1.18143>
- Ministry of Agriculture, 2016. Ethiopia Rural Land Administration System (RLAS) Manual for Maintenance of Rural Land Records. Addis Ababa.
- Musyoka, G.M., 2018. automatic delineation of small holder agricultural field boundaries using fully convolutional networks, University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, Netherlands.
- Muthukrishnan, R., Radha, M., 2011. Edge Detection Techniques For Image Segmentation. *International Journal of Computer Science and Information Technology* 3, 259–267. <https://doi.org/10.5121/ijcsit.2011.3620>
- Nichols, S., 1993. Land registration: managing information for land administration.
- Oliveira, S.A., Kaplan, F., di Lenardo, I., 2017. Machine Vision Algorithms On Cadaster Plans, in: Lewis, R., Raynor, C., Forest, D., Sinatra, M., Sinclair, S. (Eds.), 12th Annual International Conference of the Alliance of Digital Humanities Organizations, {DH} 2017, Montréal, Canada, August 8-11, 2017, Conference Abstracts. Alliance of Digital Humanities Organizations {(ADHO)}.
- Persello, C., Stein, A., 2017. Deep Fully Convolutional Networks for the Detection of Informal Settlements in VHR Images. *IEEE Geosci. Remote Sens. Lett.* 14, 2325–2329. <https://doi.org/10.1109/LGRS.2017.2763738>
- Persello, C., Tolpekin, V.A., Bergado, J.R., de By, R.A., 2019. Delineation of agricultural fields in smallholder farms from satellite images using fully convolutional networks and combinatorial grouping. *Remote Sensing of Environment* 231, 1–18. <https://doi.org/10.1016/j.rse.2019.111253>
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9351, 234–241.
- 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, 2514–2520. <https://doi.org/10.1109/36.964989>

- Shelhamer, E., Long, J., Darrell, T., 2017. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39, 640–651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- Shibeshi Gebeyehu, 2011. Cadastral Template 2.0 [WWW Document]. URL <http://cadastraltemplate.org/ethiopia.php> (accessed 4.29.22).
- Singh, S., Suresh, M., Jain, K., 2015. Land Information Extraction with Boundary Preservation for High Resolution Satellite Image. *International Journal of Computer Applications* 120, 39–43. <https://doi.org/10.5120/21243-4014>
- Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y.S.E., Dauphin, Y., Keyzers, D., Neumann, M., Cisse, M., Quinn, J., 2021. Continental-Scale Building Detection from High Resolution Satellite Imagery 1–15.
- Stöcker, C., Ho, S., Bennett, R., Koeva, M., Schmidt, C., Nkerabigwi, P., Zevenbergen, J., 2018. Towards UAV-based Land Tenure Data Acquisition in Rwanda : Needs Assessment and Technology Response, in: *FIG Congress 2018: Embracing Our Smart World Where the Continents Connect: Enhancing the Geospatial Maturity of Societies*. Istanbul, pp. 79–97.
- Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Cardoso, M.J., 2017. Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10553 LNCS, 240–248. [https://doi.org/10.1007/978-3-319-67558-9\\_28](https://doi.org/10.1007/978-3-319-67558-9_28)
- Sun, X., Zhao, W., Maretto, R. v., Persello, C., 2021. Building Polygon Extraction from Aerial Images and Digital Surface Models with a Frame Field Learning Framework. *Remote Sensing* 13, 4700. <https://doi.org/10.3390/rs13224700>
- Tuladhar, A., 1996. Spatial cadastral boundary concepts and uncertainty in parcel based information systems. *ISPRS Journal of Photogrammetry and Remote Sensing XXXI*, 890–893.
- UNECE, 1996. *Land Administration Guidelines*, United Nations. United Nations Publication, Geneva.
- UN-HABITAT, 2018. *SDG Indicator 1.4.2 Training Module: Secure Tenure Rights to Land*. Nairobi, Kenya.
- V.A. Tolpekin and A. Stein, 2013. *The core of GIScience: a systems-based approach*. University of Twente Faculty of Geo-Information and Earth Observation (ITC), Enschede, Netherlands.
- Vacca, G., 2019. Overview of open source software for close range photogrammetry. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 42, 239–245. <https://doi.org/10.5194/ISPRS-ARCHIVES-XLII-4-W14-239-2019>
- Van Asperen, P.C.M., 2014. Evaluation of innovative land tools in sub-Saharan Africa: Three cases from a peri-urban context. IOS Press. <https://doi.org/https://doi.org/10.4233/UUID:46562C4A-8E99-422F-B2E3-6949358299F6>
- Wang, S., Waldner, F., Lobell, D.B., 2022. Unlocking large-scale crop field delineation in smallholder farming systems with transfer learning and weak supervision.
- Wierzbicki, D., Matuk, O., Bielecka, E., 2021. Polish Cadastre Modernization with Remotely Extracted Buildings from High-Resolution Aerial Orthoimagery and Airborne LiDAR. *Remote Sens.* 2021, Vol. 13, Page 611 13, 611. <https://doi.org/10.3390/RS13040611>
- Wassie, Y.A., Koeva, M.N., Bennett, R.M., Lemmen, C.H.J., 2018. A procedure for semi-automated cadastral boundary feature extraction from high-resolution satellite imagery. *Journal of Spatial Science* 63, 75–92. <https://doi.org/10.1080/14498596.2017.1345667>
- Xia, X., Persello, C., Koeva, M., 2019. Deep Fully Convolutional Networks for Cadastral Boundary Detection from UAV Images. *Remote Sensing* 11, 1–14. <https://doi.org/10.3390/rs11141725>
- Zain, T., 2021. Fit-For-Purpose Land Administration In Ethiopia - Ten Years of Success, in: *Smart Surveyors for Land and Water Management*. pp. 1–14.

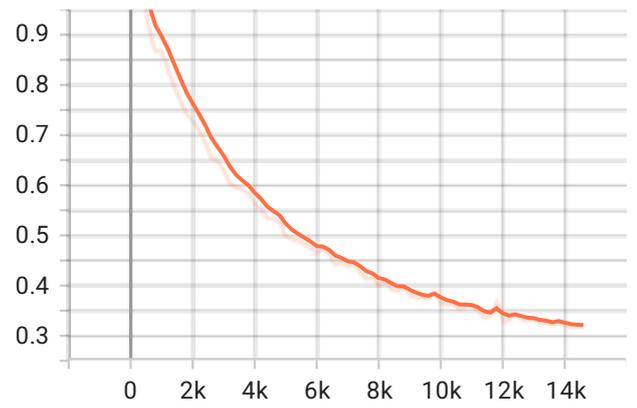
- Zevenbergen, J., Augustinus, C., Antonio, D., Bennett, R., 2013. Pro-poor land administration: Principles for recording the land rights of the underrepresented. *Land Use Policy* 31, 595–604. <https://doi.org/10.1016/j.landusepol.2012.09.005>
- Zhang, Liangpei, Zhang, Lefei, Du, B., 2016. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine* 4, 22–40. <https://doi.org/10.1109/MGRS.2016.2540798>
- Zhang, Y., Mehta, S., Caspi, A., 2021. Rethinking Semantic Segmentation Evaluation for Explainability and Model Selection. <https://doi.org/10.48550/arxiv.2101.08418>
- Zhao, W., Persello, C., Stein, A., 2021. Building outline delineation: From aerial images to polygons with an improved end-to-end learning framework. *ISPRS Journal of Photogrammetry and Remote Sensing* 175, 119–131. <https://doi.org/10.1016/j.isprsjprs.2021.02.014>
- Zorzi, S., Bittner, K., Fraundorfer, F., 2021. Machine-learned Regularization and Polygonization of Building Segmentation Masks, in: 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, pp. 3098–3105. <https://doi.org/10.1109/ICPR48806.2021.9412866>

# APPENDIX 1: LEARNING CURVES OF THE FRAME FIELD MODEL

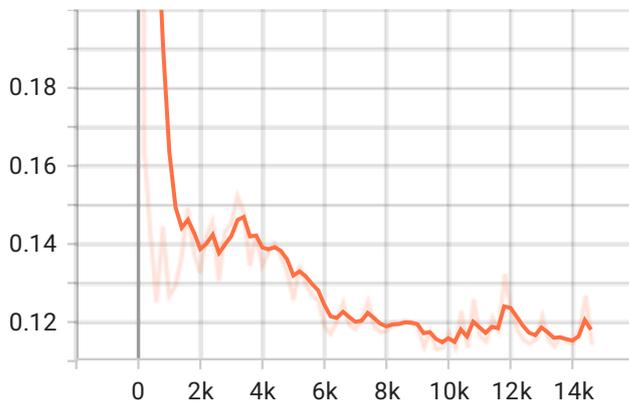




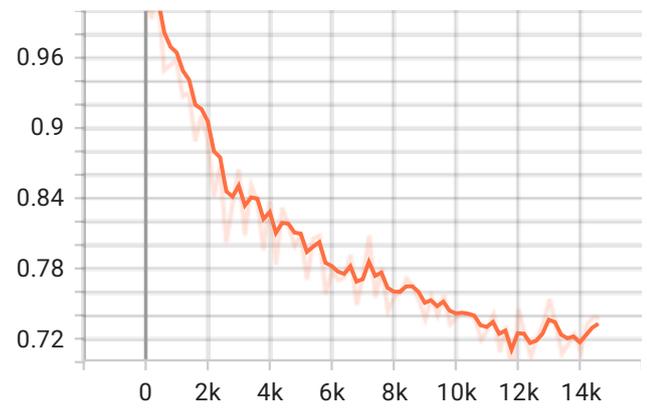
seg\_edge\_crossfield



seg\_edge\_interior



seg\_interior\_crossfield



seg