MASTER THESIS

In Order to Obtain the

PROFESSIONAL MASTER

in

GEOGRAPHIC INFORMATION SYSTEMS & DATA SCIENCE (GIS)

Presented and defended by:

Eng. HASSANIN ZOUGHAIB

Title

Enhancing Agricultural Management in Lebanon: A Comparative Study of Machine Learning Strategies for Parcel Delineation Using satellite image.

Supervisor

Dr. GHALEB FAOUR

HANINE SERHAN

Reviewers

Dr. AHMAD FAOUR

Lebanese University-Faculty of science

Acknowledgment

I want to express my sincere gratitude to my supervisors, Dr. Ghaleb Faour and Hanine Serhan, for their invaluable guidance, time, and support. I am also deeply grateful to my family, friends, and colleagues for their support and encouragement throughout this research. Their unwavering support has motivated me to push my limits and strive for success. To everyone who has been part of this journey, thank you for believing in me and inspiring me to achieve my goals.

Abstract

Detecting spatial features from satellite imagery is crucial in a variety of applications, including automated map generation, urban planning, and geographic information systems (GIS). These tasks often require accurate identification and delineation of land parcels and boundaries to ensure the precision and usability of spatial data. This research investigates applying deep learning techniques to segment and classify land parcels from satellite imagery, presenting a solution that combines accuracy with computational efficiency to advance automated map production.

The study explores various deep learning models and architectures, employing several datasets with diverse spectral band combinations, such as RGB (Red, Green, Blue) and NRG (Near Infrared, Red, Green), to evaluate their effectiveness in parcel segmentation. Each model was rigorously trained and tested to detect and delineate parcel boundaries, allowing for a comprehensive comparison of their performance across different configurations. This approach enabled the identification of key factors influencing segmentation accuracy, including the choice of spectral bands, model architecture, and training parameters.

Our experiments demonstrated that certain model configurations significantly outperformed others, achieving a superior balance between precision and recall. Notably, a specific band combination was found to enhance the model's ability to detect subtle parcel boundaries and minimize errors in segmentation. This configuration delivered highly accurate results, establishing it as the most effective setup for the given datasets and use case.

In addition to achieving high accuracy, the study highlights the advantages of deep learning techniques over traditional image processing methods. While traditional approaches often struggle with complex spatial patterns and require extensive manual intervention, deep learning models exhibit the capacity to learn intricate features and adapt to varying landscape characteristics. This advantage makes them a reliable choice for tasks requiring high precision and scalability.

Table of content

Table of Figures

List of Tables

Abbreviations

FFPLA	Fit-For-Purpose Land Administration
RS	Remote sensing
GIS	Geographic Information System
UN	United Nations
CNN	Convolutional Neural Network
FCN	Fully Convolutional Networks
RGB	Red Green Blue
NRG	Near Infrared Red Green
NRGB	Near Infrared Red Green Blue
USGS	United States Geological Survey
UAV	Unmanned Aerial Vehicles
UGV	Unmanned ground vehicles
DIP	Digital Image Procession
AI	Artificial Intelligence
AR	Augmented Reality
BDE	Boundary displacement errors
GeoTIF	Geographic Tagged Image File Format
GAN	Generative Adversarial Network
CNRS-L	National Council for Scientific Research in Lebanon
IOU	Intersection over Union
NDVI	Normalized Difference Vegetation Index
DSS	Decision Support Systems

Introduction

Agricultural management in Lebanon faces significant challenges due to the absence of comprehensive cadastral data, with approximately 40% of the territory lacking formal land registration. This thesis addresses this critical issue by employing advanced machine learning techniques to delineate agricultural parcels using multispectral high-resolution satellite imagery. The primary objective is to enhance the accuracy and efficiency of parcel delineation, thereby contributing to more effective agricultural management.

“Land parcels are the new layer of how humans own, occupy, and use the Earth, with relevance to almost any geographic question that requires question that an action” (Jerry Paffendore). A parcel is a contiguous land with defined boundaries considered a single unit for ownership, management, or legal purposes. Parcels are usually delineated based on cadastral information, including geographic coordinates, physical boundaries (such as fences, roads, or natural features), and land use, ownership, or zoning attributes. The importance of having a boundary of agriculture plays an important role in effective agricultural management which is crucial for ensuring food security, sustainable land use, and economic stability, particularly in regions where agricultural activities play a central role in the economy. In Lebanon, agriculture remains a vital sector, yet its management is significantly hindered by the absence of comprehensive cadastral data. The lack of accurate parcel delineation data presents substantial challenges, impeding land ownership clarity, agricultural planning, and resource allocation. This approach not only aids in better land management but also enhances the efficiency of agricultural practices by providing precise information on parcel boundaries. They provide accurate, accessible, and transparent information about land parcels, facilitating decision-making processes and supporting sustainable development and this impacts parcel segmentation.

The main objective is to evaluate and optimize traditional segmentation techniques for their effectiveness in delineating agricultural parcels, in addition to assessing the efficacy of integrating remote sensing with cutting-edge machine learning frameworks that incorporate innovative features for improved adaptability and precision. Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance. Special cameras collect remotely sensed images, which help researchers "sense" things about the Earth **(USGS web)**. Satellite imagery is one of the most common types of remote sensing images. It involves capturing images of the Earth's surface from satellites orbiting the planet. Satellite imagery has revolutionized the way we view and understand our planet. It allows us to monitor changes in land cover, track the movement of icebergs, and even detect wildfires from space. With advancements in technology, satellite imagery has become more accessible and high-resolution, enabling us to study the Earth in unprecedented detail. Despite the widespread use of remote sensing studies to map parcel delineation, it is still challenging to map parcel delineation due to changes in the agricultural cycle or by buying and selling lands every year. Nowadays, the manual annotation of high-resolution satellite images via visual digitization is used to determine the boundary of land delineation, which is highly dependent on the competence of individuals to analyses parcel delineation field satellite images. However, this practice was found to be costly and time-consuming, hence an automated approach is needed. The advent of high-resolution satellite imagery and advancements in computational techniques, such as machine learning, offer promising solutions to overcome these challenges. By leveraging multispectral high-resolution images, it is possible to delineate agricultural parcels with improved accuracy, even in areas where formal cadastral data is unavailable. This approach not only aids in better land management but also enhances the efficiency of agricultural practices by providing precise information on parcel boundaries. This thesis aims to address the critical issue of parcel delineation in Lebanon by employing a comparative study of traditional and advanced machine learning techniques.

Fit-For-Purpose Land Administration (FFPLA), a method increasingly adopted by organizations like UN-HABITAT and the World Bank, addresses this by utilizing modern technologies such as remote sensing and GPS to expedite the land registration process. FFPLA is designed to be flexible, affordable, and efficient, particularly in countries lacking formal cadastral systems. It aims to secure land ownership for farmers and promote stability, especially in underdeveloped regions. To support the implementation of FFPLA, the development of specialized algorithms that automate land parcel delineation is essential. These algorithms would streamline the process, ensuring accurate and timely land registration, which is critical for agricultural management and economic stability.

This thesis is focused on Zahle Bekaa Valley, especially in the plain land agriculture area. A thorough literature review is conducted to explore existing research on the use of satellite imagery in agricultural monitoring, with a particular focus on land parcel segmentation methods.

This review serves as the foundation for the research questions and hypotheses, highlighting the necessity of sophisticated computational approaches in the context of Lebanon's cadastral challenges.

The proposed methodology involves acquiring and preprocessing a comprehensive dataset of multispectral high-resolution satellite images, followed by the implementation of both traditional computational techniques and state-of-the-art machine learning models. The comparative analysis of these approaches will provide valuable insights into their effectiveness and potential for application in regions with similar cadastral challenges.

In summary, this research aims to contribute to the field of agricultural management in Lebanon by providing practical, data-driven solutions for parcel delineation, ultimately supporting more efficient and sustainable agricultural practices. The findings of this thesis are to explore and develop various deep learning models for the precise delineation of agricultural parcels using satellite imagery with 50cm of spatial resolution. By implementing multiple models, the study aims to identify and generate high-quality geospatial data that are not only spatially accurate but also reliable for agricultural decision-making processes. The models will be rigorously evaluated based on their predictive accuracy, to determine the most effective approach for parcel segmentation. The findings of this evaluation will support the selection of the most suitable model, capable of enhancing precision agriculture practices by providing accurate parcel boundaries, which are crucial for tasks like crop monitoring, resource allocation and land management.

Chapter 1: Literature Review

The delineation of agricultural parcels presents a significant challenge in land management, particularly in areas with insufficient cadastral data. Over time, various approaches have been developed to tackle this issue, ranging from traditional image segmentation techniques to contemporary machine learning algorithms. This literature review aims to critically assess these methods, with a particular emphasis on their relevance to the Lebanese context, where around 40% of the land lacks formal registration.

Satellite imagery has emerged as a fundamental tool for automating numerous tasks such as map production, Geographic Information Systems (GIS), agricultural monitoring, and urban planning. Unlike conventional imagery, satellite data provides more structured and consistent spatial information, making it particularly suitable for applications like road extraction, building footprint detection, and land cover classification. Many of these tasks depend on semantic segmentation models to accurately extract valuable information from the imagery.

The potential applications of satellite-derived data are extensive, encompassing fields such as urban planning, environmental monitoring, GIS, and fleet management, among others. While the use of a limited, labeled dataset of satellite imagery can constrain a model's ability to accurately identify spatial features, numerous studies have effectively utilized semantic segmentation models to recognize various elements, including roads, buildings, and vegetation.

Before exploring related studies and their critical evaluation, we will define and clarify key concepts from remote sensing, machine learning, deep learning, computer vision, and semantic segmentation, focusing on their applications in agriculture and parcel delineation.

1.1. Remote Sensing

1.1.1. Overview

Remote sensing (RS) is the acquisition of information about an object or phenomenon from distance. This involves an instrument or a sensor mounted on a platform, such as a satellite, an aircraft, an UAV/UGV, or a probe. The sensor typically measures the electromagnetic radiation that is either reflected or emitted by the target. The type of information accessible from remote sensing depends on the specific properties of the instrument and its platform. These properties include: satellite or biography, UAV/UGV motion plan, field sensor position and orientation, active or passive sensing, detector array and optical lens characteristics, as well as storage capabilities.[7]

The processed image is interpreted visually or electronically or digitally to extract the information about the illuminated target. Remote sensing systems which measure reflected energy are called passive sensors, which can be used only to detect energy in the presence of naturally occurring energy. This can take place only during the time when the sun is illuminating the earth. An active sensor provides its own energy source for illumination. The sensors emit radiation which is directed towards the target to be investigated; these sensors obtain the information regardless of the time of day. In order to capture the earth's surface, the sensors must be placed in a proper platform. Before it was ground-based and aircrafts platforms, nowadays satellite near-polar orbits platform provides a great contribution to remote sensing imagery (Demir et al. ,2018). More over Multispectral satellite sensor provides digital raster images, that allow us to apply Digital Image Processing (DIP) techniques to develop thematic maps of landuse/landcover classes which are essential in many remote sensing applications like forestry, agriculture, environmental studies, weather forecasting, ocean studies, archeological studies etc.

RS has become an indispensable tool in modern agricultural management, providing valuable data for monitoring and assessing crop health, land usage, and environmental conditions. By using satellite imagery and other airborne sensors, remote sensing enables the collection of large-scale spatial data without physical contact, making it ideal for observing agricultural parcels over time. Remote sensing technology offers critical insights into factors like vegetation indices, soil moisture, and land use patterns. In agricultural management, remote sensing aids in estimating crop yields, monitoring water stress, and ensuring sustainable farming practices. Over the past decade, it has evolved to support precision agriculture, reducing the reliance on manual data collection by automating the analysis of large datasets (Ashraf et al., 2023) [1]

1.1.2. Different Types of Imageries

Satellite imagery plays a crucial role in RS, with different types of satellite sensors providing varied resolutions and spectral bands that can be utilized for agricultural analysis. In the past, moderate-resolution imagery, such as Sentinel-2 (10-meter resolution), was commonly used for regional-scale studies due to its accessibility and multispectral capabilities. However, higher-resolution imagery, such as that provided by WorldView-3, PlanetScope, or the 50 cm resolution satellite imagery used in this thesis, has opened new avenues for detailed parcel delineation and crop management.

High-resolution imagery is particularly important for precision agriculture, where the ability to delineate small and fragmented agricultural parcels is crucial. Studies such as García-Pedrero et al.'s work on agglomerative segmentation for agricultural parcel delineation have shown that using high-resolution imagery can improve the accuracy of parcel boundary detection, which is a key factor in estimating subsidies and managing land resources. In the following section, we will delve deeper into the various types of imagery resolutions, clarifying what is meant by high-resolution and its significance in RS applications.

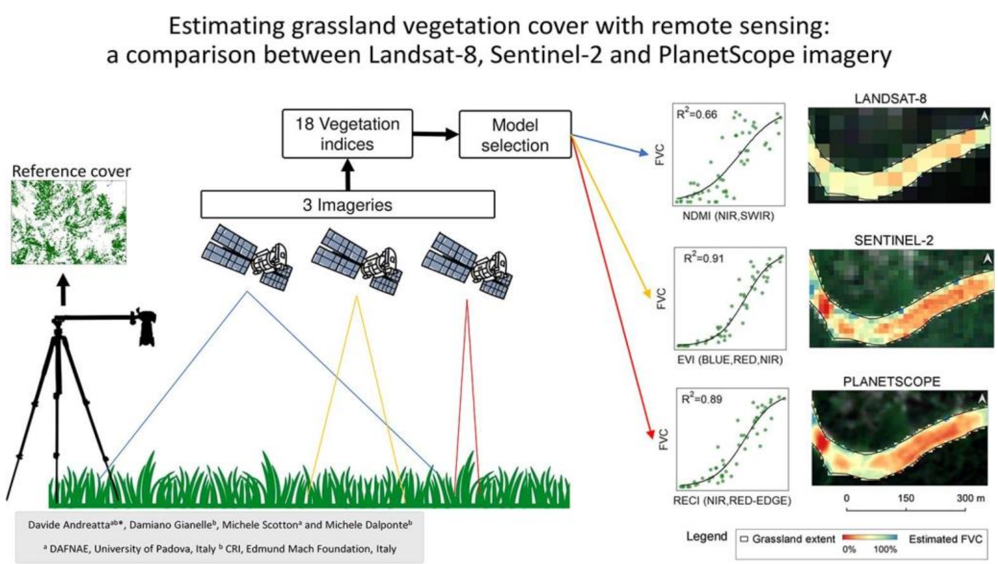


Figure 1: Comparison between Landsat-8, Sentinel-2 and Planet

1.1.3. Image Resolution and Band Combination

In RS, several types of resolution affect the quality and applicability of satellite imagery for agricultural analysis. Understanding the nuances of each type is crucial for selecting the appropriate data for tasks such as parcel delineation, crop monitoring, and yield estimation. These include spatial, temporal, spectral, and radiometric resolutions:

- **Spatial Resolution** refers to the size of the smallest object that can be detected by the sensor, typically measured in meters or centimeters. High spatial resolution imagery, such as the 50 cm imagery used in this thesis, allows for the detection of fine details, making it ideal for identifying small, fragmented agricultural parcels and intricate field patterns. The finer the spatial resolution, the more precise the delineation of boundaries and monitoring of crop health.[12]

- **Temporal Resolution** defines the frequency at which a satellite revisits the same location on Earth. For agricultural applications, high temporal resolution is essential for tracking changes over time, such as crop growth, phenology, or detecting early signs of stress or disease. Satellites like Sentinel-2 offer a revisit time of 5 days, while commercial satellites such as PlanetScope provide near-daily observations, allowing for frequent monitoring of crop conditions.[12]

- **Spectral Resolution** indicates the number and width of spectral bands captured by the sensor. Multispectral imagery typically includes visible bands (red, green, blue) and additional bands like near-infrared (NIR), which are critical for agricultural analysis. NIR is particularly useful for calculating vegetation indices such as NDVI, which provides insights into plant health, biomass, and water stress. Some satellites, such as WorldView-3, offer even higher spectral resolution with more specialized bands.

Images	Description
Multispectral	A multispectral image consists of several bands of data. The bands in the multi-spectral image are less than 30. This depends on the application and it is always less than 30 bands
Super-spectral	Contemporary satellite sensors are skilled, concerning the capturing of images at several higher wavelength bands. For instance, numerous satellites are composed of thirty-six spectral bands, managing the region's wavelength ranges from the near-infrared, visible, and shortwave infrared to the thermal infrared. The band bandwidth is restricted, allowing exact spectral aspects of the targets to be acquired via the sensor.
Hyperspectral	The hyperspectral image has hundreds or more adjacent spectral bands with a three-dimensional image cube. The compression of dimensions is utilized in hyperspectral image examination for decreasing data volume and redundancy.

Table 1: Various Remote Sensing Images Spectral Resolutions

- **Radiometric Resolution** refers to the sensor's ability to distinguish between different levels of brightness or reflectance. This is measured in bits, with higher radiometric resolution providing more detailed information about the intensity of light reflected from the surface. In agriculture, higher radiometric resolution helps in distinguishing subtle differences in vegetation health, soil moisture, and other key factors.

Combining these resolutions allows for more precise monitoring of agricultural land. For example, using high spatial resolution imagery with multiple spectral bands enables accurate delineation of crop boundaries while simultaneously providing valuable insights into crop health through spectral analysis. High temporal resolution ensures that farmers and decision-makers can continuously monitor their fields throughout the growing season, responding to changes in near real-time.

1.2. Geographic Information System

1.2.1. Overview

A Geographic Information System (GIS) is a powerful tool that enables the collection, management, analysis, and visualization of spatial data. By integrating data from various sources, including remote sensing, GIS supports the mapping and monitoring of land use, crop health, and other agricultural variables. GIS has revolutionized agricultural management, allowing for the efficient allocation of resources, optimized irrigation strategies, and precision farming practices. The spatial data layers in GIS can include information on soil properties, crop types, and water availability, providing farmers and decision-makers with the tools to make informed decisions.

1.2.2. Practical Use of GIS in Agriculture

GIS is widely used in agriculture for applications such as mapping field boundaries, monitoring crop health, and managing irrigation systems. By combining remote sensing data with GIS, it is possible to create detailed maps that provide insights into soil conditions, crop yield potential, and pest infestations. For example, farmers can use GIS to visualize crop health variations across a field and apply targeted interventions, such as localized fertilization or irrigation adjustments.

Moreover, GIS is instrumental in the delineation of agricultural parcels, allowing for the accurate mapping of field boundaries, which is essential for tasks like land tenure and subsidy allocation. Recent advancements in GIS technology, including integration with machine learning algorithms, have enhanced its ability to analyze large datasets and provide real-time updates to farmers and policymakers.

1.3. Traditional ways of segmentation

Before the rise of machine learning and deep learning algorithms, traditional image segmentation methods were widely used in remote sensing and agricultural applications. Image segmentation originally started from Digital Image Processing coupled with optimization algorithms. These primitive algorithms made use of methods like region growing and snake's algorithm where they set up initial regions and the algorithm compared pixel values to gain an idea of the segment map.[13] They primarily relied on pixel-based approaches and manual techniques to delineate boundaries between different land parcels or agricultural fields.

Figure 2: Image and its segmentation mask

Although they provided a foundation for early segmentation work, these techniques faced limitations in handling complex, heterogeneous landscapes, especially in fragmented agricultural settings. Algorithms that took a global view of the input image came much later on with the following methods:

- **Thresholding**, one of the simplest segmentation techniques, involves selecting a range of pixel values to differentiate between regions in an image. In agricultural applications, this technique could be used to separate crops from background areas such as soil or water by selecting thresholds based on pixel intensity or vegetation indices. However, thresholding is often insufficient for complex images with varying lighting conditions, plant growth stages, or diverse vegetation types. [12]

- **Edge Detection** methods aim to identify the boundaries of objects by detecting changes in pixel intensity. Techniques such as the Canny or Sobel edge detectors were commonly applied to detect the edges of fields, roads, or other infrastructure in satellite images. In agricultural parcel delineation, edge detection helped identify field boundaries, but it often struggled with noisy data and intraclass variability. Additionally, it required careful parameter tuning and manual intervention to refine the results.
- **Region-based** methods, such as region-growing algorithms, group adjacent pixels based on similarity in color or texture to form larger segments. This method works well when the objects of interest (e.g., fields or plots) are relatively uniform in appearance. However, in agricultural landscapes where parcels can vary significantly in size, shape, and crop type, region-based methods can over-segment the image, producing too many small regions that do not correspond to actual fields.

- **Watershed** algorithm treats an image as a topographic surface, where pixel intensities represent elevation. Its "floods" the image from low-intensity areas to high-intensity areas, identifying ridgelines that correspond to object boundaries. While the watershed algorithm can successfully segment distinct objects, it is highly sensitive to noise and may produce an excessive number of regions, especially in agricultural landscapes where interparcel variability is high.

Traditional segmentation methods, while foundational, faced several limitations when applied to agricultural imagery:

- Sensitivity to noise: Many traditional methods struggle with variations in lighting, shadow, and crop growth, leading to inaccurate boundaries.
- Over-segmentation: These methods often break up large fields into smaller, irrelevant segments due to variability within parcels.
- Manual intervention: Most traditional approaches require significant manual tuning and post-processing to achieve acceptable results, which is time-consuming and subjective.

These limitations highlighted the need for more advanced techniques that could handle complex, high-resolution imagery more efficiently and accurately. This paved the way for the introduction of machine learning and, later, deep learning models, which could automate and enhance the segmentation process with greater precision.

1.4. Machine learning for image analysis

1.4.1. Overview

Machine learning is a subset of artificial intelligence (AI) that enables computers to learn from data and improve their performance on a specific task without being explicitly programmed. Unlike traditional programming, where rules are hardcoded, machine learning algorithms identify patterns and make decisions or predictions based on the data they're trained on. These algorithms leverage statistical patterns and inference to discover underlying relationships within data and enhance their accuracy over time as they are exposed to more information. The goal of machine learning is to empower computers to learn from experience and make reliable and accurate predictions or decisions.

Geospatial applications require specialized tools to collect, process, and present geographical data. These tools are invaluable for urban planning, environmental monitoring, resource management, and other fields. Their primary advantage lies in providing valuable insights for decision-making and resource allocation, contributing to sustainable societal development. Spatial data refers to information tied to specific locations on Earth, such as satellite imagery, maps, and geospatial data. Machine learning can revolutionize spatial data detection from satellite imagery by automating the identification and analysis of patterns.

This includes tasks like land cover classification, land use change detection, and monitoring natural disasters or environmental shifts. By feeding vast amounts of satellite imagery to machine learning algorithms, these algorithms can learn to recognize patterns and features within the images. This capability can be used to develop predictive models that identify areas at risk of natural disasters or environmental hazards, monitor ecosystem health, and support urban planning.

1.4.2. Supervised and Unsupervised learning

CLASSIFICATION METHODS

There are several approaches and methods that are associated with satellite image classification. But most widely satellite images are classified into two main categories as shown in the Figure 3 i.e. Pixel-based as well as Object-based. The pixel-based techniques are further divided into unsupervised and supervised techniques. A brief about various classification methods are tabulated in the Table 2.[10]

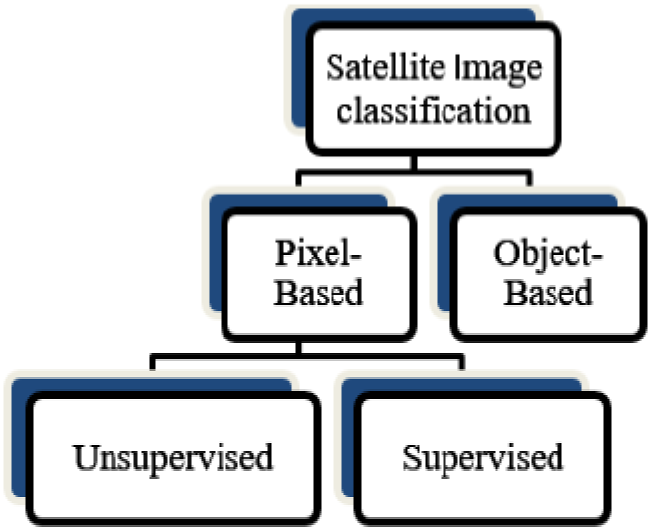


Figure 3: Satellite image classification techniques

Methods	**Description**
Pixel-based classification	Pixel-based classifications are based on the grey value of pixels and for the classification purpose only spectral information is used. These are considered as the least unit that depicts some image. This technique utilizes the statistics of reflectance for particular pixels. It assembles pixels to express land cover features. The land coverage has to be forested, metropolitan, agricultural and another features variety. The classification of the pixel is further categorized into unsupervised as well as supervised classification

Unsupervised Classification	In this, pixels are integrated according to the reflectance properties. These groups are individually called as 'Clusters'. The analyst classifies the varied clusters to produce and the bands to utilize. The analyst recognizes the clusters with the classes of land coverage classes. Later, analyst allocates significant labels to the clusters and delivers properly satellite image. It is frequently the case in which several clusters stand for a distinct land cover class. The analyst integrates the clusters into a land cover category. This classification method is usually employed while no sample locations exist. K-means and ISODATA are the techniques used for unsupervised classification of the satellite images
Supervised Classification	In the Supervised classification, an input is required from the analyst. The input of the analysts is termed as 'training set'. The sample of training is considered to be an important aspect of the methods of supervised classification and the accuracy of these methods vastly on the samples employed for the purpose of training. The classification is based on the spectral signatures in the training set. Each class is demonstrated on the basis of what it is similar to the most in the sets of training. Foremost classifiers for supervised classification are Minimum-distance, parallelepiped and maximum likelihood. The algorithm primarily separates the pixels from each other based on the training samples that denote a site on the ground.
Object-Based Classification	It is different from pixel-based classification approach as it works on the group of pixels instead of direct pixels. In Object-Based image classification, some image is interpreted not only for single pixel but is also valid in significant image objects and their common relationships.

	Object-based information extraction is relies upon spectrum character, geometry as well as structure evidence. This method offers actually innovative data and can simply accessible. Object-Based Classification has two main stages: i. Image Segmentation to generate a segmented image ii. Classification of the segmented image.

Table 2: Classification Methods

- **Supervised Classification:**

The supervised classification is the essential tool used for extracting quantitative information from remotely sensed image data. Using this method, the analyst has available sufficient known pixels to generate representative parameters for each class of interest.

This step is called training. Once trained, the classifier is then used to attach labels to all the image pixels according to the trained parameters. The quality of a supervised classification depends on the quality of the training sites, The most commonly used supervised classification is maximum likelihood classification.[26]

- **Unsupervised Classification:**

Pixels are grouped based on the reflectance properties of pixels. These groupings are called clusters. The user identifies the number of clusters and bands to be generated. With this information, the image classification tool generates clusters. There are different image clustering algorithms such as K-means and Expectation Maximization. The unsupervised classification technique is commonly used when no sample sites exist.

Criteria	Categories	Characteristics	Classifiers examples
Usage of training samples	Supervised	Providing land cover classes. Appropriate reference data is given and can be utilized as training samples. The signatures produced are utilized for training the classifiers for classifying the spectral data in the thematic map.	More likelihood Less distance The classifier of the decision tree
	Un-supervised	The algorithm of clustering is utilized for portioning the spectral image into different spectral classes on the basis of statistical data from the image. There is no usage of prior classes definition. It is the responsibility of the analyst to label and merge the spectral	ISODATA K-mean clustering algorithm

		classes in desired classes.	
Utilization of different parameters like Mean Vector/Covariance matrix	Parametric classifiers	The assumption of the Gaussian distribution. The generation of parameters from training samples. The noisy results are used when the landscape is complex. It is tough to combine the spatial, contextual and ancillary data into the classification process.	More Likelihood LDA (Linear Discriminant Analysis
	Non-Parametric classifiers	The data is not assumed for the requirement. It doesn't need statistical parameters for calculating the class separation and is appropriate for integration of non-	ANN (Artificial Neural Network) Evidential reasoning SVM Decision tree classifier Expert System

		remote sensing data in the classification process.	
Pixel information type	Per-pixel classifiers	Signature is developed by traditional classifiers by integrating the training set spectra of the pixels via feature. The signature has the contribution of each material within the pixels of the training sets without combined pixel problems.	Classifiers like more likelihood Less distance ANN SVM Decision tree
	Sub-pixel classifiers	The spectral value for the pixel is taken as linear/ non-linear with the integration .	Fuzzy set classifiers. Spectral mixture analysis Sub pixel classifiers
The output is specific for land cover class	Hard classification	Every pixel should be allocated to the unique class	More likelihood Less distance
		The estimation of the area may result in errors from data of coarse spatial	ANN SVM Decision-tree

		resolution because of the problem of mixed pixels.	
	Soft classification	A similarity degree is provided for every class.	The classifiers of fuzzy set.
		It gives more information and more precise results	The analysis of spectral mixture Sub-pixel classifiers.

Table 3: Image Classification Method Taxonomy

1.4.3. Application in Agriculture

Semantic segmentation has become a transformative tool in agriculture by enabling more precise and data-driven decision-making processes that enhance both productivity and sustainability. One of its primary applications is in precision agriculture, where it allows for detailed monitoring of crop growth and health across vast agricultural fields. By leveraging high-resolution satellite or drone images, semantic segmentation models can differentiate between crop types and stages of growth, helping farmers make informed decisions about irrigation, fertilization, and harvesting schedules. For example, crops in different growth stages can be identified and segmented, allowing targeted interventions that optimize inputs like water and nutrients based on the plant's current needs.

Beyond monitoring, targeted weed and pest management is another crucial area where segmentation excels. Weeds compete with crops for nutrients, light, and water, significantly impacting crop yield. Traditional weed management techniques often involve spraying herbicides uniformly across fields, which can be wasteful and harmful to the environment. Semantic segmentation offers a solution by accurately distinguishing weeds from crops, allowing for more precise herbicide application, reducing the amount of chemicals used, and lowering costs. This precision can extend to pest detection, where early signs of infestation can be spotted and treated efficiently, further protecting crop yields.

Another important benefit of semantic segmentation is in yield prediction. By using historical data and current imagery, segmentation models can identify patterns in plant growth and health, which can be linked to expected yields. This data-driven approach provides more accurate predictions than traditional methods, allowing farmers to anticipate production levels and plan their resources and logistics accordingly. For example, areas with underperforming crops can be identified early, enabling farmers to apply corrective measures before significant losses occur.

In terms of resource management, segmentation helps optimize irrigation practices. By analyzing moisture levels across different zones in a field, semantic segmentation enables precise water allocation. This prevents over-irrigation in certain areas while ensuring that under-watered regions receive adequate attention, promoting better water conservation. Similarly, soil health assessment benefits from this technology by enabling the identification of soil conditions that might impact crop growth, such as nutrient deficiencies, compaction, or erosion. Farmers can then take corrective actions, such as applying fertilizers or adjusting their planting strategies to improve soil conditions.

Moreover, semantic segmentation is instrumental in disease detection, an area where early intervention is critical. Many plant diseases manifest as subtle changes in leaf color or texture, which can be difficult to detect with the naked eye. Segmentation models, particularly when combined with multispectral or hyperspectral imaging, can identify these early signs of disease at the pixel level, allowing for rapid intervention before the disease spreads widely. This capability not only saves crops but also reduces the need for extensive pesticide use, promoting more sustainable agricultural practices.

In the realm of automation, segmentation supports the development of robotic and drone technologies used for tasks like automated harvesting and precision spraying. By accurately segmenting crops, roads, and other features in the field, these systems can operate autonomously with minimal human intervention. For instance, drones equipped with semantic segmentation algorithms can apply pesticides or fertilizers only to the areas that need them, reducing waste and enhancing operational efficiency.

Finally, segmentation aids in land-use planning and crop rotation, allowing farmers to delineate specific parcels of land and classify them according to their use, whether for active crop production, fallow periods, or infrastructure like roads and irrigation systems. This helps optimize land management by ensuring that different areas of the farm are used effectively, contributing to long-term soil health and productivity. The accurate mapping of land parcels also supports the effective rotation of crops, which is essential for maintaining soil fertility and preventing pest build-up.

Semantic segmentation in agriculture not only improves the precision and sustainability of farming operations but also empowers farmers with actionable insights to optimize their use of resources. By reducing waste, enhancing crop health, and promoting early detection of potential problems, this technology plays a pivotal role in advancing modern agricultural practices and ensuring food security.

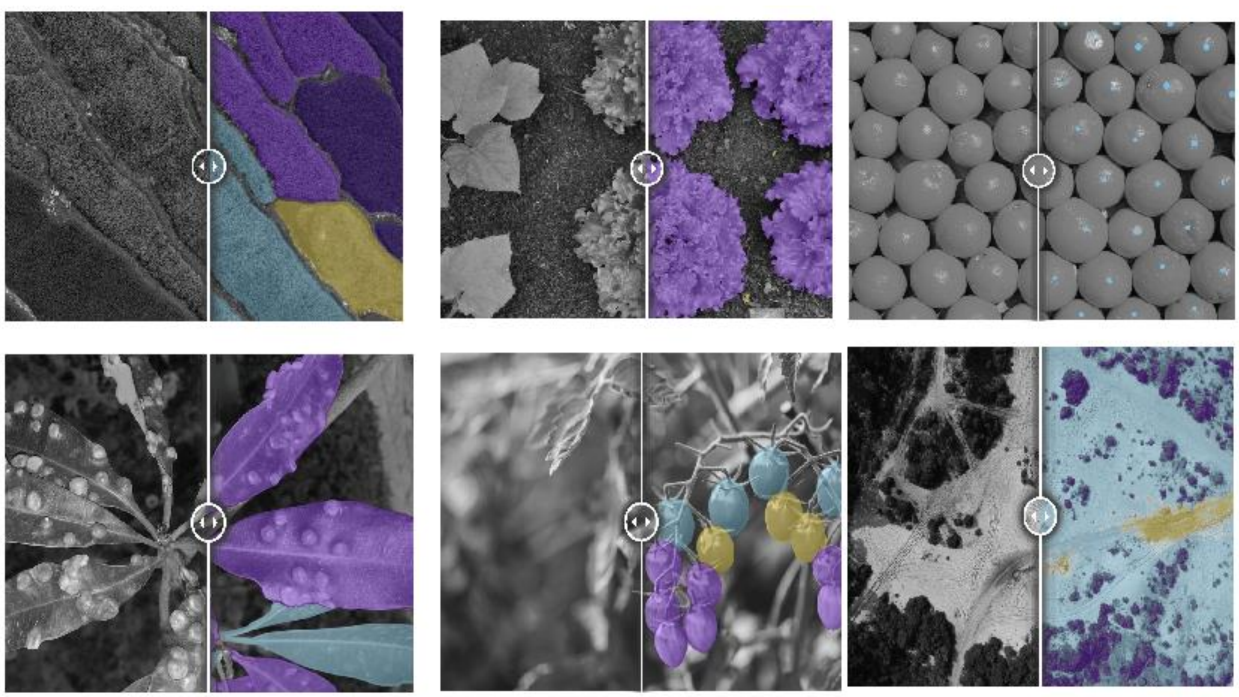


Figure 4: Some applications of Deep learning in agriculture use.

1.5. Deep Learning

1.5.1. Overview

Deep learning, a machine learning technique inspired by the human brain, statistics, and applied mathematics, has gained significant popularity and practicality in recent years. Advances in computer hardware, larger datasets, and training techniques for deeper networks have fueled its growth. While offering great potential, deep learning also presents challenges and opportunities for further development.

As a contemporary supervised learning method, deep learning provides a powerful framework. By incorporating more layers and units within a layer, deep networks can represent increasingly complex functions. This approach excels at tasks involving mapping input vectors to output vectors, which humans can often perform effortlessly. However, deep learning requires extensive models and labeled training data. Complex tasks that cannot be easily described as vector-to-vector mappings or require substantial human reasoning remain beyond its current capabilities.

Remote sensing is one application of deep learning in spatial data analysis. Deep learning algorithms can analyze satellite imagery to identify and classify various features on Earth's surface, such as land cover types, vegetation density, and urban areas. This information enables monitoring environmental changes, such as deforestation or urbanization, and informs decision-making in fields like urban planning, agriculture, and environmental conservation.[16]

While AI is a broad field, machine learning is a specific application that allows machines to learn from data without explicit programming. Machine learning often employs simpler methods like decision trees or linear regression to extract knowledge from data, whereas deep learning utilizes more advanced methods found in artificial neural networks.

Deep learning requires less human intervention, as it can automatically extract features from a dataset. In contrast, simpler machine learning techniques often necessitate manual feature identification and classifier selection by engineers. Deep learning can learn from its own errors, while machine learning typically requires human intervention. Deep learning also demands significantly more data and computational power than machine learning. While machine learning can often be performed on servers with CPUs, deep learning frequently requires more powerful chips like GPUs.

1.5.2. Computer Vision

Computer Vision is a field of artificial intelligence that enables computer systems to extract information from images or videos, by using digital images from cameras, videos, or other sensors. It aims to emulate human vision capabilities, allowing computers to process and analyze visual information in a manner similar to the human brain.

1. **Early Beginnings (1960s-1970s):**

Computer vision as a formal discipline began in the 1960s and 1970s, with early work focusing on basic image processing and pattern recognition. The initial research was primarily theoretical and aimed at understanding the fundamental problems of extracting meaningful information from images. Notable early milestones included edge detection: techniques like the Sobel and Canny edge detectors were developed to identify boundaries within images and object Recognition: Early algorithms focused on identifying simple shapes and patterns.[4]

2. Growth and Development (1980s-1990s):

The 1980s and 1990s saw significant advancements in computer vision due to increased computational power and more sophisticated algorithms. Key developments included:

- Feature Extraction: Techniques for detecting and describing features in images, such as corners and textures, became more refined.
- Machine Learning: The introduction of machine learning methods provided new approaches for training systems to recognize objects and patterns.
- 3D Vision: Research expanded to include 3D modeling and reconstruction from 2D images, enabling more complex analyses of visual data.

3. Modern Era (2000s-Present):

The 2000s marked a new era of rapid progress in computer vision, driven by advancements in deep learning and neural networks. Major factors contributing to this progress include:

- Deep Learning: The development and application of convolutional neural networks (CNNs) revolutionized image classification and object detection, achieving state-of-the-art performance on various benchmarks.
- Big Data and Computing Power: The availability of large datasets and powerful GPUs enabled training more complex models, leading to breakthroughs in tasks such as facial recognition and autonomous driving.
- Integration with Other Technologies: Computer vision systems began to integrate with other AI technologies, such as natural language processing and robotics, creating more sophisticated and versatile applications.

There is different application of computer vision like Image Classification, Object Detection, Image Segmentation, Face Recognition, Motion Tracking and Augmented Reality (AR).

1.5.3. Semantic Segmentation

As opposed to image classification, in which an entire image is classified according to a label, image segmentation involves detecting and classifying individual objects within the image. Additionally, segmentation differs from object detection in that it works at the pixel level to determine the contours of objects within an image.

Semantic segmentation is a task for partitioning an image into segments to be able to detect objects from the images by assigning a semantic label to each pixel of an image. It is used to identify the boundaries of objects in images [2].

Deep learning has exhibited remarkable accuracy in computer vision tasks and holds tremendous potential for efficiently processing vast amounts of earth observation satellite image data in automated workflows. [18].

The field of computer vision has been a thriving research area for deep learning applications, primarily due to the inherent complexity of vision, which humans and animals effortlessly perform but poses significant challenges for computers [20]. Computer vision is a highly expensive discipline that encompasses diverse image-processing techniques and a multitude of applications. Its scope spans from emulating human visual capabilities, like facial recognition, to pioneering novel visual abilities. Common benchmark tasks for evaluating deep learning algorithms in computer vision include optical character recognition and object recognition. [26].

One of the most challenges in the history of computer vision is Semantic Segmentation because it requires the algorithm to not only detect objects in an image but also to precisely segment them into their individual parts. Unlike object detection, where the goal is to identify the location of an object in an image, semantic segmentation requires pixel-level labeling of each object in the image. Furthermore, the need for high precision and accuracy in semantic segmentation makes it particularly challenging. Even small errors in the segmentation of an object can have significant consequences in downstream applications, such as autonomous driving, where a misclassified object could result in a collision.

Image segmentation is the process of partitioning an image into multiple segments or regions to simplify its analysis. The goal is to segment the image into regions that are meaningful and easier to analyze. Each segment typically corresponds to distinct objects or areas of interest within the image. This process helps in understanding the image content more effectively, as it breaks down complex images into simpler, more manageable parts.

Also, there are different type of segmentation:

1. **Semantic Segmentation** that assigns a class label to every pixel in the image, grouping pixels that belong to the same object or class into the same segment the aim is to identify objects within an image, but it does not differentiate between different instances of the same object class and it can be used in a street scene, semantic segmentation would label all pixels belonging to cars as "car", all pixels belonging to pedestrians as "pedestrian", without distinguishing between individual cars or pedestrians.

2. **Instance Segmentation** goes a step further by not only classifying pixels into categories but also distinguishing between different instances of the same object class. It provides more detailed information by identifying and segmenting each object instance separately and it used in a crowded scene with multiple people, instance segmentation would label each person individually, creating distinct segments for each individual, even if they are of the same class (e.g., all labeled as "person").

3. **Panoptic Segmentation** combines elements of both semantic and instance segmentation. It provides segmentation at the object level, including both class labels and instance differentiation. This type of segmentation provides a comprehensive understanding of the image by segmenting objects while also recognizing their instances it can be used in a photo with several overlapping objects like various fruits in a bowl, panoptic segmentation would identify and segment each fruit separately while also classifying them correctly.

Image segmentation is a fundamental aspect of computer vision with diverse techniques and applications. By breaking down images into meaningful segments, it enables more detailed and accurate analysis of visual information. As technology advances, particularly with the rise of deep learning, segmentation methods continue to improve, expanding their capabilities and applications across various domains.

1.6 Related Work

The segmentation of agricultural parcels and crop monitoring through remote sensing has seen significant advancements with the advent of deep learning models. Traditional approaches like thresholding and edge detection have paved the way for machine learning and deep learning methods that can handle complex agricultural landscapes with greater precision. One of the most notable works in this domain is [12], where a machine learning approach was used to automate the delineation of agricultural parcels. This study introduced a methodology based on agglomerative segmentation, using super pixels as the foundational unit for image segmentation. The results showed significant improvement in delineating fragmented agricultural parcels from high-resolution satellite imagery. Additionally, there have been several works on the integration of deep learning with remote sensing data to monitor crop health and field boundaries.

In this section, we compare several notable studies that employ different deep learning architectures, datasets, and techniques for agricultural segmentation. The objective is to assess the effectiveness of these models, highlight their strengths and limitations, and illustrate how they have contributed to the evolving landscape of remote sensing applications in agriculture. By evaluating these works, we can better understand the unique contributions of each model, particularly focusing on the U-Net architecture, which has become one of the most widely used models in the field.

Despite the advancements, challenges remain in accurately delineating parcels in areas with high spatial heterogeneity, where deep learning models must be adapted to local conditions and available data. Nonetheless, the general trend in the literature supports the conclusion that U-Net and other deep learning models are highly effective for agricultural image segmentation and are becoming the standard in precision agriculture applications.

Several studies have explored deep learning models for agricultural segmentation, crop classification, and field boundary delineation. Each approach utilizes different methodologies, datasets, and evaluation metrics, highlighting the strengths and challenges of applying machine learning and deep learning techniques to remote sensing imagery. The comparison of these models illustrates the diversity of approaches to agricultural segmentation, each tailored to specific data types and problem domains.

- **U-Net Model for Agricultural Plot Delineation:**

The U-Net architecture is one of the most widely used models for agricultural plot segmentation due to its efficiency in semantic segmentation tasks. In [12], García-Pedrero et al. (2019) utilized a U-Net-based approach to automatically delineate agricultural parcels from high-resolution orthophotos. Their model outperformed traditional edge-detection-based methods, such as the gPb-UCM algorithm, by significantly reducing boundary displacement errors (BDE). The U-Net model's encoder-decoder architecture allowed it to capture both global and local features, crucial for accurately delineating agricultural boundaries in fragmented landscapes.

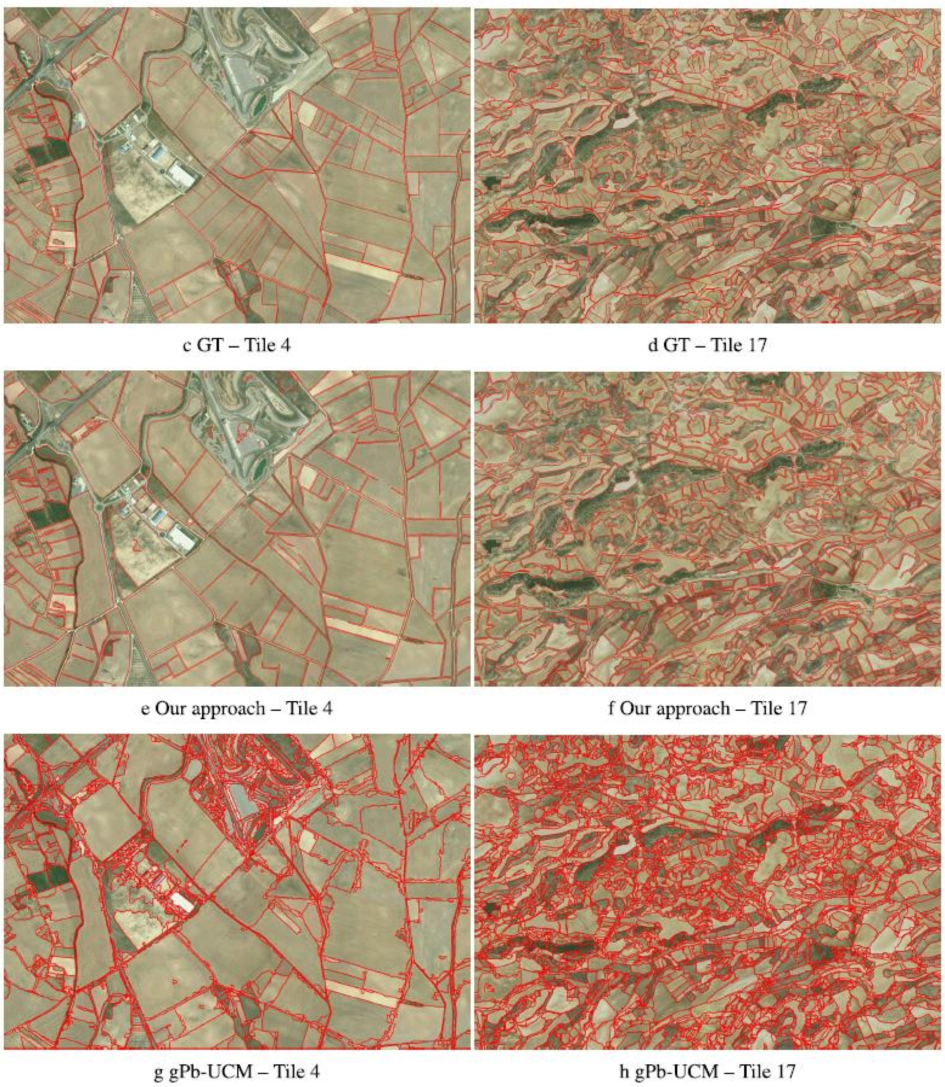


Figure 5: Result for different models.

The study also introduced the use of the Land Parcel Identification System (LPIS) data to train the U-Net model. The LPIS data provided labeled ground truth for the delineation task, improving the model's performance in identifying the correct boundaries. By comparing their U-Net-based method with gPb-UCM, the authors showed that the CNN-based model achieved higher accuracy, especially in heterogeneous landscapes where traditional methods failed.

- **Farm Parcel Delineation Using Spatio-temporal Convolutional Networks:**

The main purpose of this project was to make boundary for farm parcels instead of using ‘theodolites, total stations, and GPS’ by applying an AI model using deep learning methods.

In recent years, deep learning has become very popular in computer vision tasks due to its incredible success, so for this purpose they use a pixel classification method using U-net architecture to apply such model, the use a sentinel 2 image with 3 bands RGB for agriculture area in France (2017).

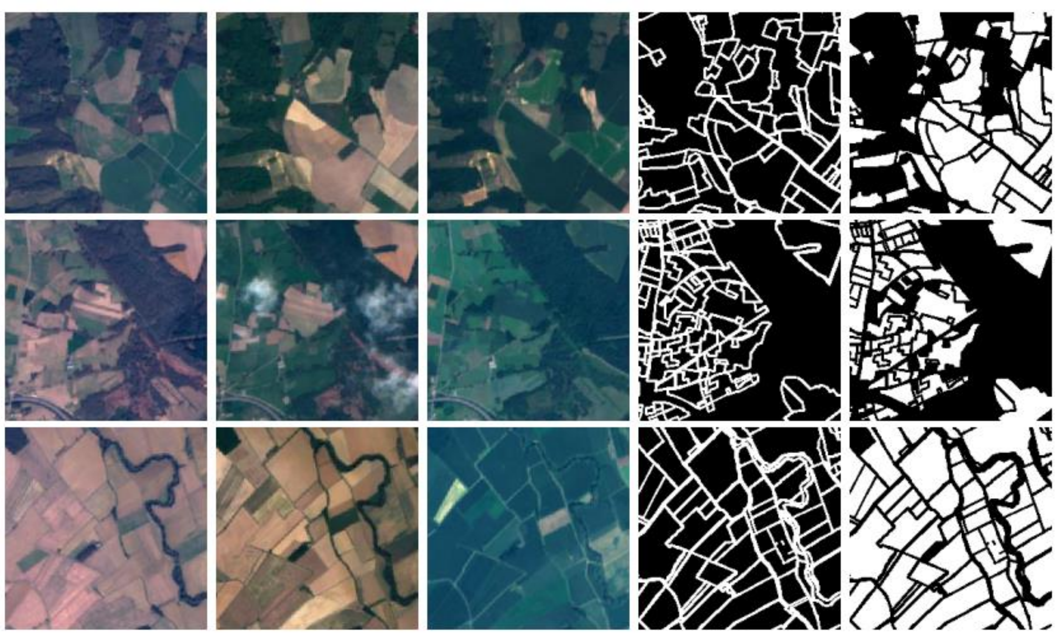


Figure 6 : images and predicted tiles for parcels segmentation.

In particular, the trained variants of the U-Net model on the Sentinel-2 images given the corresponding area/boundary masks. They showed that the proposed Spatio-temporal U-Net achieves 83% Dice score.

- **Extraction of Parcel Boundary from UAV Images Using Deep Learning Techniques:**

The article presents a deep learning-based method for extracting cadastral parcel boundaries from (unmanned aerial vehicle) UAV images, aiming to simplify and automate the traditionally labor-intensive and time-consuming process of boundary delineation. Using a U-Net-based CNN architecture, they successfully developed a model that generates accurate parcel boundaries with minimal human intervention. The model was trained on a dataset prepared from UAV images and corresponding vector boundaries, achieving satisfactory results in predicting parcel boundaries. However, the authors acknowledge the need for further improvements, particularly to handle varied data sources and resolutions, and suggest exploring additional methods such as data augmentation, GANs, and automation of data preparation for even more robust results. This approach has promising applications for ongoing cadastral boundary updates, potentially transforming fieldwork by providing efficient and automated boundary delineation.

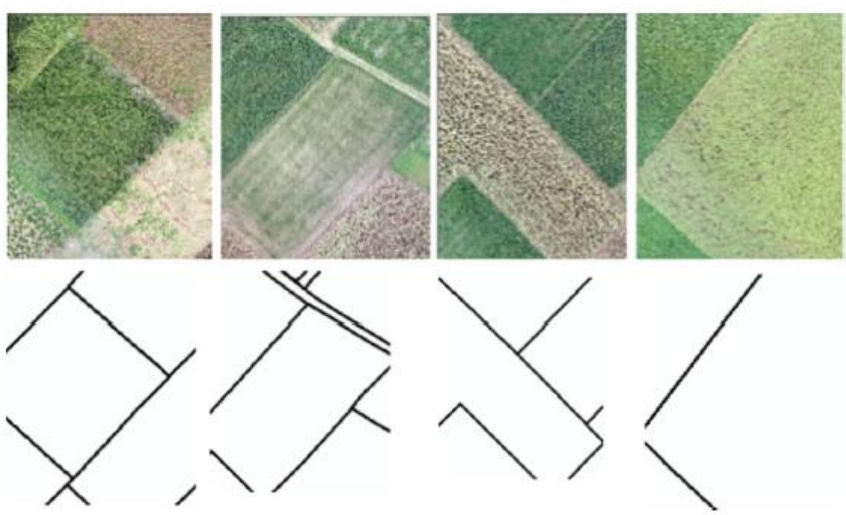


Figure 7:images and labels tiles for parcels segmentation.

Conclusion

U-Net is a convolutional neural network that was developed for image segmentation. The network is based on a fully convolutional neural network whose architecture was modified and extended to work with fewer training images and to yield more precise segmentation, as many projects and related works use this type of deep learning model U-Net seems to be a strong choice for pixel classification.

Parcel segmentation requires accurately delineating parcel boundaries, which can be complex and irregular. U-Net’s skip connections allow it to capture high-resolution, fine-grained details, making it well-suited to identifying the often subtle boundaries of parcels in imagery also parcels in imagery can vary significantly in shape and size, from small, irregular patches to large blocks. U-Net’s encoder-decoder structure, combined with skip connections, allows it to capture both large contextual features and fine details, making it well-suited for this variability in parcel characteristics. Parcel segmentation often relies on remote sensing or satellite data, which may include multiple spectral bands (like RGB, and near-infrared). U-Net has been shown to perform well with multi-channel inputs, allowing it to learn from various spectral features that enhance parcel detection.

Overall, U-Net’s ability to capture detailed boundaries, handle diverse parcel characteristics, and work effectively with multi-spectral imagery makes it an excellent choice for precise and reliable parcel segmentation. Also besides using u-net architecture, there are different backbones that can be used, backbone is a pre-trained convolutional neural network (CNN) that serves as the feature extraction layer for various tasks, like object detection, classification, or segmentation. It forms the core architecture that learns to identify and extract important features from images, such as edges, textures, shapes, and patterns, which later help in distinguishing objects or regions in an image.

In the context of a U-Net model for semantic segmentation, the backbone provides a strong foundation for recognizing spatial features across different regions of the image. The backbone serves as the encoder portion, which is responsible for learning and extracting important features from the input image. Selecting an appropriate backbone for U-Net can significantly impact the model's performance, efficiency, and ability to generalize.

Chapter 2: Methodology

This chapter outlines the methodology used for developing a deep learning model to delineate agricultural parcels using satellite imagery. As shown in the following flowchart in figure 3, the methodology is divided into several key steps:

1. **Data acquisition:** Collect relevant satellite imagery data for area of interest from different source and it depend purpose or accuracy needed to build such model.
2. **Data preparation:** Conduct necessary preprocessing, such as cloud masking, image enhancement, and radiometric correction, georeferencing, orthorectification.
3. **Data labeling:** refers to the process of annotating data with specific tags or classifications to teach a model what different objects or regions represent.
4. **Data training:** It’s a process for teaching the AI model to recognize patterns and make predictions based on labeled data.
5. **Data evaluation:** evaluate and measure the performance of the trained model to ensure its reliability.
6. **Model integration:** Embedding the trained model into a practical application or platform for use.

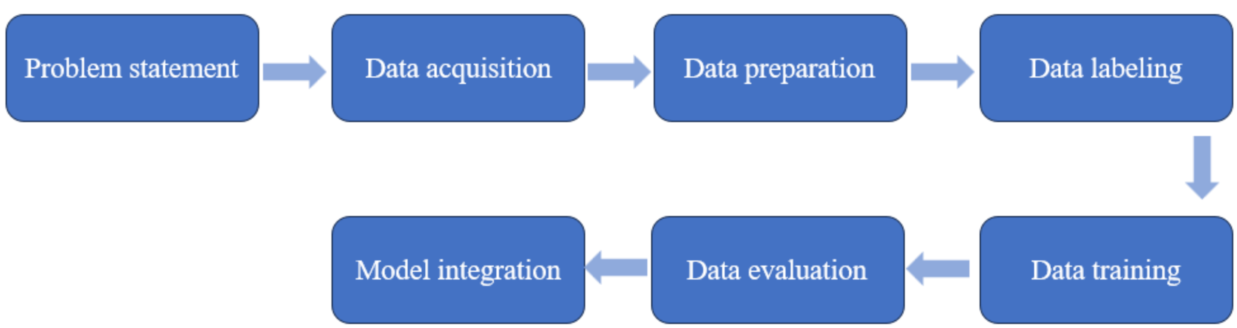


Figure 8 : Methodology flowchart

In the subsequent sections, each of these phases is examined in depth to provide a comprehensive understanding of the methodologies employed. The following sections delve into the specific processes of all phases highlighting their significance in developing an effective deep learning model for agricultural parcel delineation.

Data Collection and Preparation:

In studies focused on agricultural monitoring and parcel delineation, high-resolution satellite imagery is essential for capturing detailed land features. Satellite imagery provides data at different spatial and spectral resolutions, depending on the satellite and sensor used, and typically undergoes preprocessing to enhance its quality and usability. Common preprocessing steps include georeferencing to align the imagery with a specific coordinate system, orthorectification to correct geometric distortions due to terrain, radiometric correction to adjust pixel intensity for more accurate reflectance values, and contrast stretching to improve the visual quality of the images.

These steps are especially important when multiple images or temporal data are used, ensuring spatial consistency and enhancing feature detection capabilities.

For this study, we focused on the Zahle - Bekaa region as our study area, selected due to its significant agricultural activities. We utilized imagery from the WorldView-4 satellite, known for its very high-resolution capabilities, which are widely used by commercial, governmental, and international organizations.

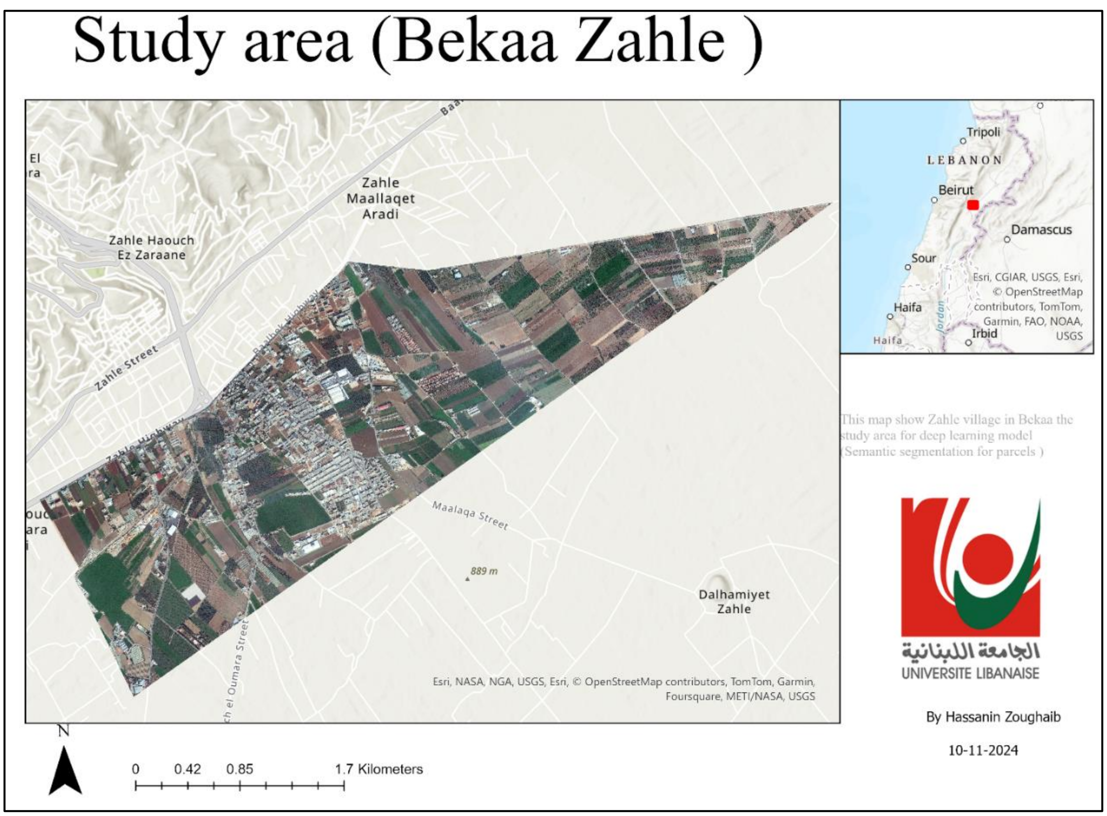


Figure 9: Study area map.

Raster Information

Columns	12952
Rows	8068
Number of Bands	4
Cell Size X	0.500021000000004
Cell Size Y	0.5000210000000281
Uncompressed Size	398.62 MB
Format	GRID Stack 7.x
Source Type	Generic
Pixel Type	unsigned char
Pixel Depth	8 Bit
NoData Value	0, 0, 0, 0
Colormap	absent
Pyramids	levels: 3, resampling: Nearest Neighbor
Compression	None
Mensuration Capabilities	Basic

Band Metadata

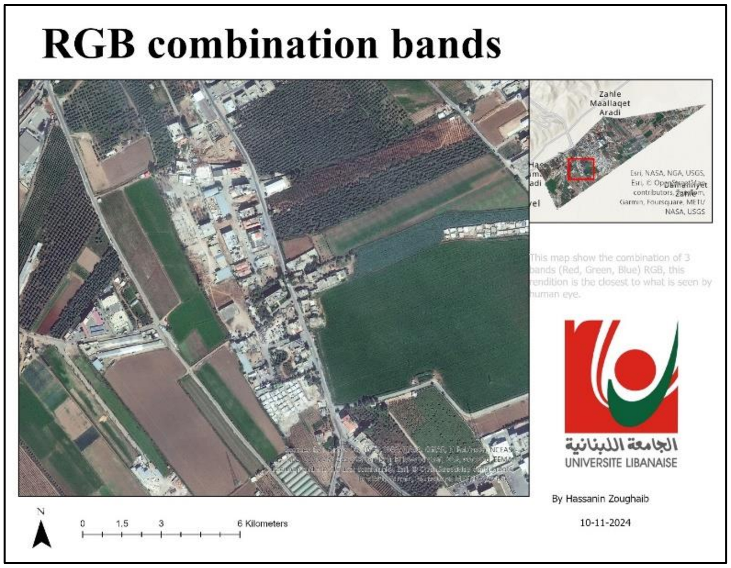
- infraredc1
- infraredc2
- infraredc3
- infraredc4

Figure 10: Raster information for satellite image (Metadata)

The WorldView-4 imagery has a spatial resolution of 0.5 meters and includes four spectral bands: Near-Infrared (N), Red (R), Green (G), and Blue (B), with an 8-bit unsigned depth for each pixel, providing detailed color information essential for distinguishing agricultural features. The data is projected in the UTM WGS 84 Zone 36N coordinate system and was delivered in GeoTIFF format, already orthorectified and cloud-free, ensuring high spatial accuracy and eliminating the need for further geometric adjustments.

The imagery utilized in this research was annotated as part of a collaborative project with the National Council for Scientific Research in Lebanon (CNRS-L). Since its establishment in 1962, CNRS-L is serving the scientific community in Lebanon covering all scientific disciplines. Its main objective is to encourage scientific research and support human resources development along the general scientific policies adopted by the government. CNRS-L is committed to keep the scientific community in Lebanon connected with advances achieved worldwide, at the same time dedicate its resources to meet local development objectives.

As a result, the imagery was immediately ready for use in our workflow without additional preprocessing. This enabled us to initiate the labeling process directly within ArcGIS Pro, where training polygons representing agricultural parcels were manually delineated.



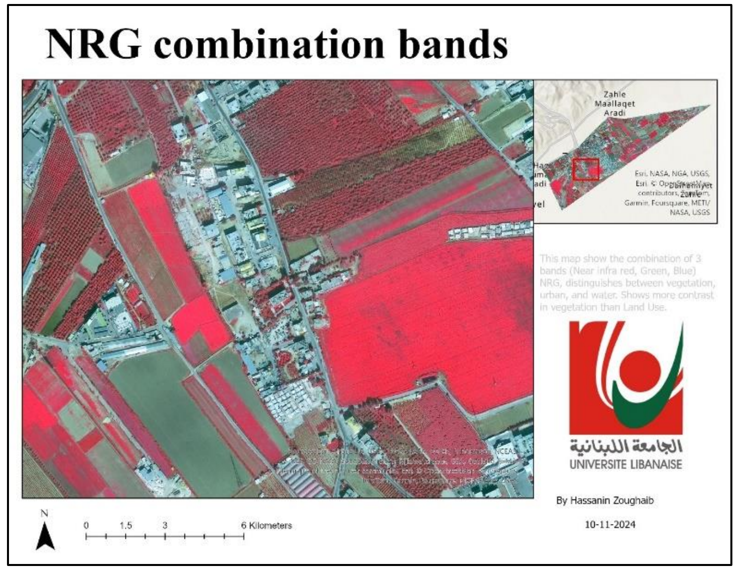


Figure 11: RGB combination bands **Figure 12: NRG combination bands**

Labeling is a crucial step in the workflow, as the model learns to identify features based on the labeled training data. Accurate labeling enables the model to distinguish between parcel and non-parcel areas, ensuring that the model produces reliable and precise segmented maps. In other words, well-labeled data is essential for the model to generate meaningful and accurate results.

For this study, labeling was performed using ArcGIS Pro 3.3. We utilized the Edit Tool to manually delineate agricultural parcels for training data in parcel segmentation. In the Zahle - Bekaa study area, approximately 750 agricultural parcels were labeled, covering a total area of around 16,000 m². These parcels exhibit a wide range of colors, shapes, and sizes, reflecting the diversity of crop types and parcel boundaries. Each parcel polygon was meticulously drawn to represent ground truth boundaries, based on visual cues such as color variations in the land and contextual features (e.g., roads, water canals, buildings, rivers, and walls). This detailed labeling serves as the foundation for training the model, enabling it to recognize and accurately delineate agricultural parcels.

Upon completing the labeling phase, the labeled parcels were exported using the Export Training Data for Deep Learning tool in ArcGIS Pro. This tool extracts imagery and corresponding labels from the GIS environment and formats them into datasets compatible with popular deep learning frameworks. The export process generated three distinct datasets, each featuring different band combinations to test their effectiveness in model training. These cases were:

- Case 1: RGB band combination
- Case 2: NRG band combination
- Case 3: NRGB band combination

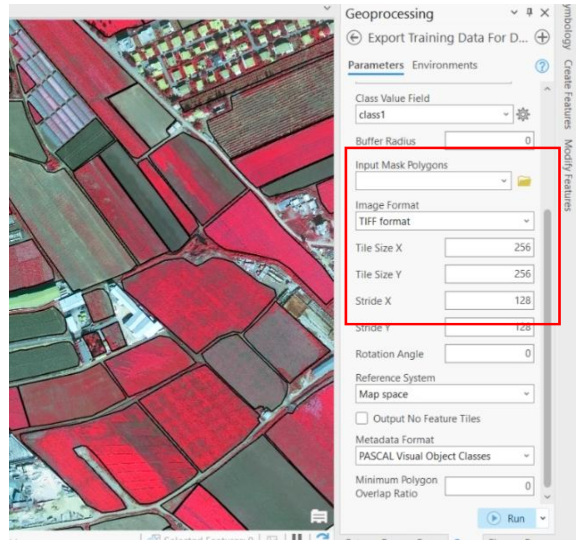
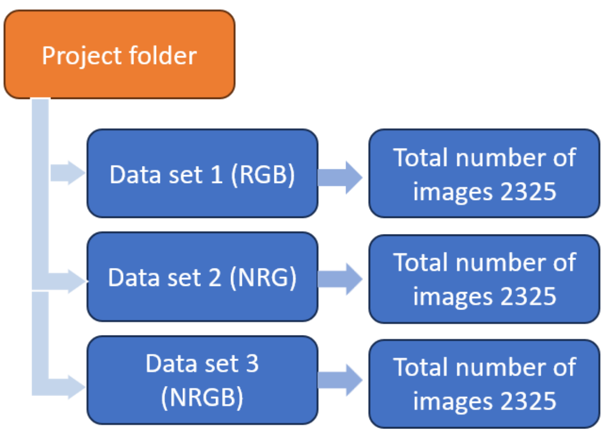


Figure 13: Export Training Data for Deep learning tool.



For each case, the tool produced a total of 2,325 image-mask pairs, with each image and mask tile having a resolution of 256x256 pixels and a stride of 128 pixels. These datasets provide a robust basis for evaluating the models' performance across various spectral inputs.

Figure 14: Different data sets.

In this project, a total of 15 experiments were conducted using the U-Net model, each with a different backbone architecture, including ResNet-34, ResNet-50, Inception-ResNet, and EfficientNet-B4. These experiments were designed to evaluate the impact of various backbone architectures on model performance, allowing us to identify the most effective configuration for agricultural parcel segmentation. The model training and evaluation processes were implemented using Jupyter Notebook, with TensorFlow and Keras as the primary deep learning frameworks.

Custom Python code was developed to train the model on each dataset variant, facilitating a structured comparison of model performance across different band combinations and architectural choices.

In the following sections, each aspect of the code and implementation strategy will be discussed in detail, covering the data preparation, model configuration, training processes, and evaluation metrics used to determine the optimal model setup for this task.

2.1 Model implementation

2.1.1 Data Augmentation techniques

Data augmentation is technique used in machine learning and computer vision to artificially expand the size of a training dataset by creating modified versions of existing data. This process helps improve the generalizability and performance of a model, especially when the available dataset is limited.

Data augmentation is a way we can reduce overfitting on models, where we increase the amount of training data using information only in our training data. The field of data augmentation is not new, and in fact, various data augmentation techniques have been applied to specific problems. The main techniques fall under the category of data warping, which is an approach which seeks to directly augment the input data to the model in data space.[28]

Many deep learning frameworks, such as PyTorch, Keras, and TensorFlow provide functions for augmenting data, principally image datasets. In image-based tasks, data augmentation typically involves applying various transformations to images to create new samples. Common data augmentation techniques include: rotation, Flipping, Scaling and Cropping, Brightness, Contrast, and Color Adjustments, Adding Noise.

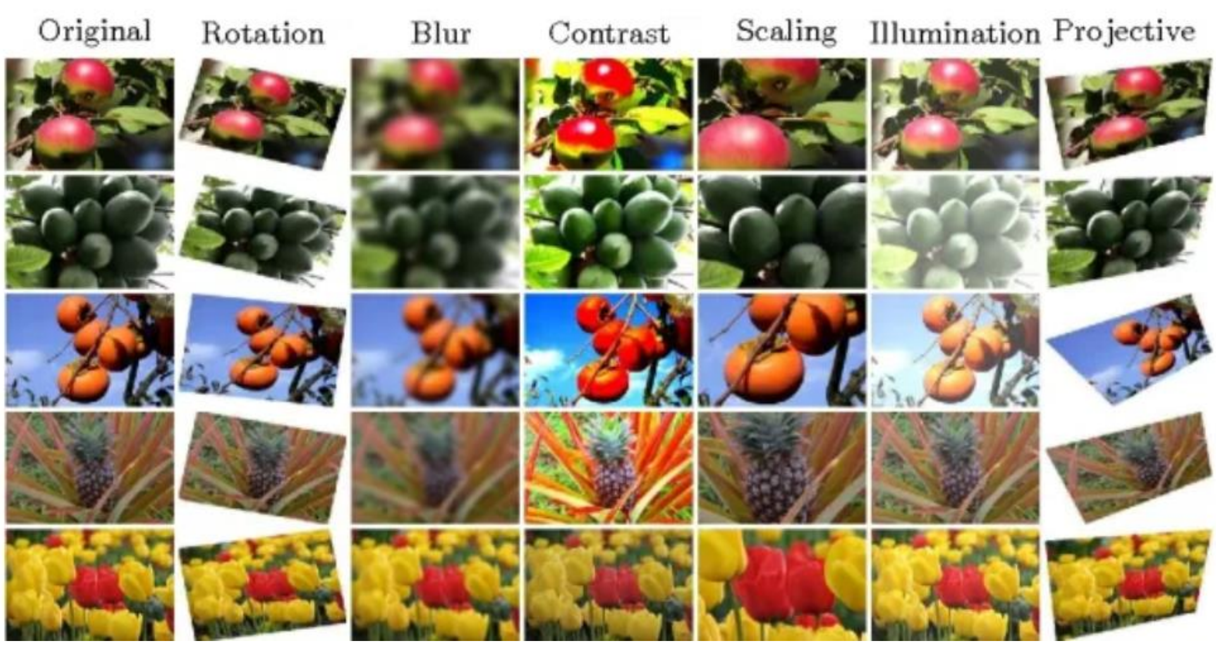


Figure 15: Augmentation types.

In this project, data augmentation plays a vital role in enhancing the robustness and performance of the model, particularly given the variations in color, shape, and size among agricultural parcels. By applying a range of transformations, we can artificially expand the training dataset, making the model more adaptable to diverse conditions and reducing the risk of overfitting. For our implementation, we used the Keras **Image Data Generator** class to perform data augmentation.

```
def get_data_generators(train_images, train_masks, val_images, val_masks):
    train_datagen = ImageDataGenerator(horizontal_flip=True,
    vertical_flip=True,
    rotation_range=90,
    zoom_range=0.2,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    fill_mode='nearest')
```

Figure 16: Augmentation code using different types.

This configuration applies several transformations:

- **Horizontal and Vertical Flipping**: Randomly flips images along both axes, introducing variety in the orientation of parcels.
- **Rotation**: Rotates images by up to 90 degrees, helping the model recognize parcels regardless of their alignment.
- **Zoom**: Randomly zooms into images by up to 20%, allowing the model to handle variability in parcel sizes.
- **Width and Height Shifts**: Shifts images by up to 20% along the width and height, simulating positional variations.
- **Shear Transformation**: Applies a shear transformation, creating slight angular distortions that mimic real-world changes in parcel shape.
- **Fill Mode**: When parts of the image are shifted or rotated beyond the original boundaries, the fill_mode='nearest' setting fills these regions with the nearest pixel values to preserve continuity.

2.1.2 U-Net Architecture for Semantic Segmentation

Olaf Ronneberger and his team introduced U-Net in 2015 as a new approach to image segmentation, particularly for medical images. It outperformed the traditional sliding window method by using fewer images and data augmentation.

The sliding window technique, while effective for localization, has two main limitations: it creates a lot of redundant information due to overlapping patches, and the training process is slow. These issues make it impractical for many tasks, U-Net addresses these problems by providing a more efficient and effective way to segment images.

The **U-Net model** is a well-known architecture in deep learning for semantic segmentation tasks, particularly effective in medical imaging and remote sensing applications. Developed initially for biomedical image segmentation, U-Net has become popular in tasks where precise delineation of objects or regions is needed, such as parcel detection in satellite imagery.

U-Net Architecture gets its name from its architecture. The "U" shaped model comprises convolutional layers and two networks. First is the encoder, which is followed by the decoder. With the U-Net, we can solve the above two questions of segmentation: **"**what" and "where.[3]

The architecture of U-Net is characterized by its **encoder-decoder structure**:

- **Encoder (Contracting Path)**: This part resembles a typical convolutional network, progressively downsampling the input image through convolutional layers and pooling. As the encoder compresses the spatial dimensions, it captures high-level contextual information.
- **Decoder (Expanding Path)**: The decoder upsamples the compressed features back to the original image dimensions, gradually recovering spatial details through a series of transposed convolutions.
- **Skip Connections**: One of U-Net's key innovations is the use of skip connections between corresponding layers in the encoder and decoder. These connections help retain fine-grained details by transferring high-resolution features from the contracting path to the expanding path, which improves the model's ability to accurately localize objects.

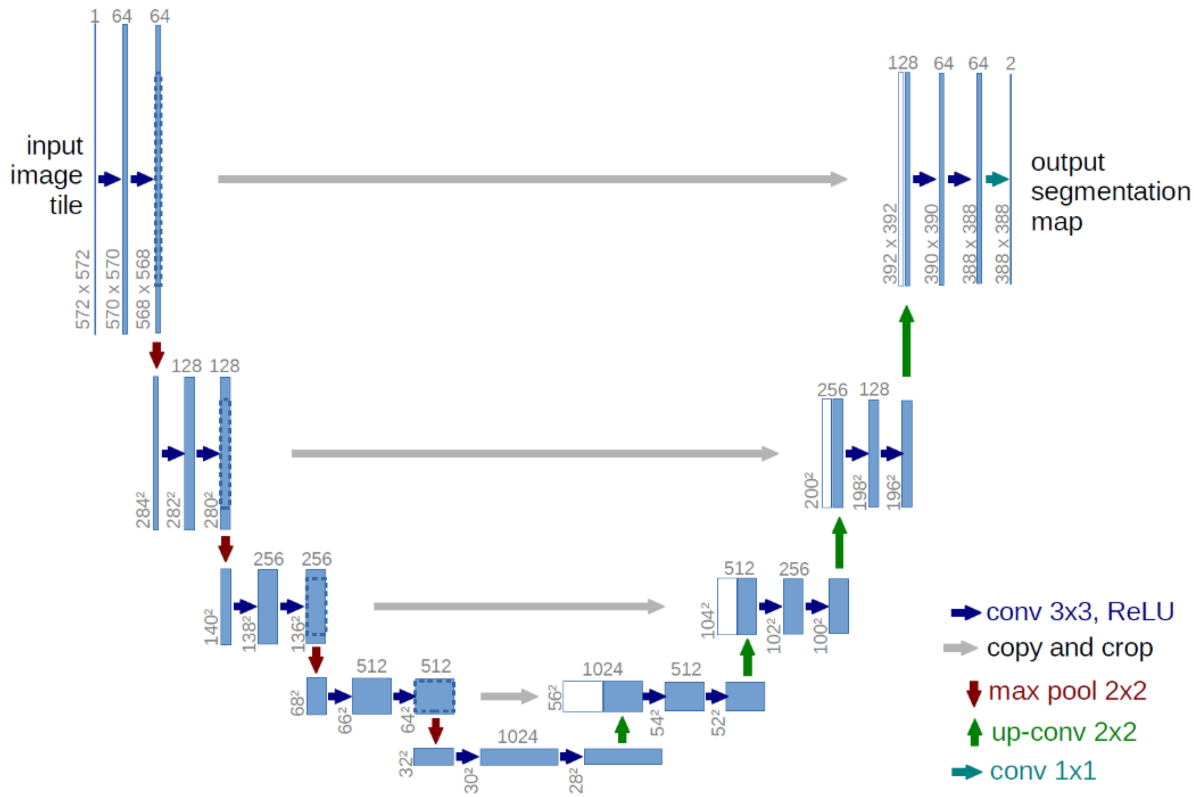


Figure 17: U-Net architecture.

2.1.3 Importance of Backbones in U-Net and Backbone Selection

In the context of deep learning, a **backbone** is a pre-trained convolutional neural network (CNN) that serves as a feature extractor in larger architectures like U-Net. By using a pre-trained backbone, we leverage its ability to detect basic patterns (e.g., edges, textures) and more complex features, significantly improving the segmentation model's performance, especially when data is limited.

For this project, we experimented with multiple backbones to optimize U-Net's performance for agricultural parcel segmentation, each has unique characteristics:

- **ResNet-34 and ResNet-50**: Known for their simplicity and efficiency, ResNet models introduce skip connections within each block, which help prevent the vanishing gradient problem in deep networks. ResNet backbones capture hierarchical features effectively, making them well-suited for tasks requiring high-level abstractions. Pre-trained on

ImageNet, these backbones have 34 and 50 layers, respectively, with ResNet-50 being deeper and capturing more complex features.

- **Inception-ResNet**: Combines the Inception architecture's multi-scale feature extraction with ResNet's residual connections, enabling it to capture patterns at various scales and achieve a good balance between accuracy and computational efficiency. This hybrid model incorporates both Inception modules and residual connections, making it adept at learning multi-scale features.

- **EfficientNet-B4**: Developed by Google, EfficientNet uses a compound scaling technique that optimally balances network depth, width, and resolution. This backbone achieves high performance with fewer parameters, making it effective for complex datasets with limited computational resources. EfficientNet-B4 is part of a series of models that trade off depth, width, and resolution for optimal accuracy. The B4 variant is moderate in size, balancing between computational cost and feature richness.
- **VGG16**: Although simpler and more computationally intensive than others, VGG16 is effective at capturing deep features and is still widely used due to its straightforward architecture. A 16-layer network that is simpler but highly effective at hierarchical feature extraction, VGG16 remains useful in cases where interpretability and simplicity are prioritized.

These backbones were chosen because they offer a diverse set of properties, allowing us to explore various trade-offs between computational complexity and feature representation capacity. By testing different backbones, we evaluated how each one impacts U-Net's ability to accurately segment agricultural parcels based on different image features.

2.1.4 Data Splitting

In machine learning, dividing the dataset into training, validation, and test sets is a foundational practice that underpins model development, ensuring that the model's performance is both robust and generalizable. This segmentation serves multiple purposes, each integral to the development of a model that not only learns from the data provided but also generalizes effectively to unseen data.

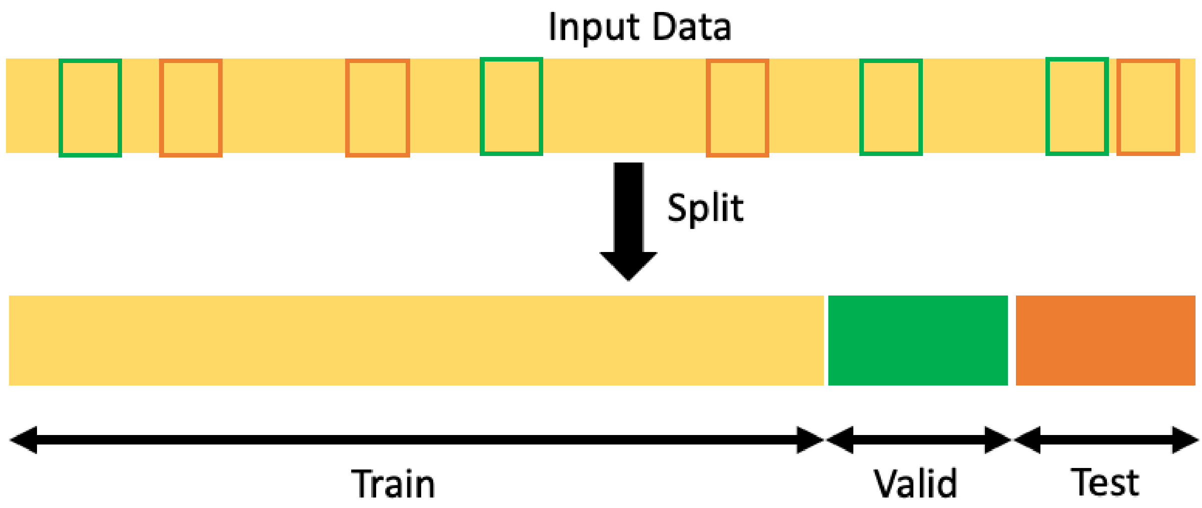


Figure 18: Data splitting.

The training set, typically comprising the largest portion of the data (in this case, 70%), is solely used for learning the underlying patterns and relationships within the data. During this phase, the model adjusts its internal parameters through optimization techniques to reduce error on the training data. However, solely optimizing on this subset can lead to a phenomenon known as overfitting, where the model becomes overly attuned to the specific characteristics of the training data, at the expense of its ability to perform well on new, unseen data. Overfitting results in a model that may exhibit high accuracy during training but performs poorly in real-world applications where data can differ in subtle ways.

To counteract overfitting and promote generalization, a validation set, representing 15% of the data, is introduced. The validation set serves as an independent dataset that the model has not encountered during training, and its role is to monitor the model's performance during training. By evaluating the model on this held-out set, it becomes possible to detect overfitting early and make adjustments. For instance, if the model performs well on the training set but poorly on the validation set, it signals that the model may be too complex and has memorized training specifics rather than general patterns. Hyperparameter tuning—adjusting model parameters such as the learning rate, regularization strength, or network architecture—relies heavily on feedback from the validation set to identify configurations that improve model performance while avoiding overfitting.

Finally, the test set, which also constitutes 15% of the dataset, plays a distinct and crucial role in assessing the model's generalization capability. Unlike the training and validation sets, which inform model building and adjustment, the test set remains untouched until the final evaluation phase. This dataset mimics real-world conditions where the model will encounter completely new data. The use of an isolated test set provides an unbiased evaluation of the model's ability to generalize to unseen data. Since the model and hyperparameters have not been adjusted based on the test set, its results represent a fair and realistic measure of how the model is expected to perform in a production environment.

Thus, the 70-15-15 split ratio strikes an essential balance. It offers a sufficient amount of data for training, which allows the model to learn and generalize effectively. Meanwhile, the validation and test sets provide reliable benchmarks for performance monitoring and final evaluation, respectively, ensuring that the model performs well across multiple phases of assessment. This balanced approach not only safeguards the model against overfitting but also instills confidence in its predictive performance when applied to real-world data, a critical factor in the successful deployment of machine learning models.

2.1.5 Model Implementation

The U-Net model code is an AI-based pipeline for semantic segmentation, specifically designed to classify pixels in images, such as parcel detection in geographic images. It is organized into functions that handle various stages of the workflow, from setting up the environment, preparing data, defining the model, training, and evaluating, to visualizing results. The implementation of our U-Net model was carried out using the **TensorFlow** and **Keras** libraries, which offer robust tools for building and training deep learning models. We organized our code into modular scripts that handle different stages of the segmentation workflow, including data preparation, model building, training, and evaluation, each contributing to an efficient and reproducible training pipeline.

To begin, we set up the necessary environment by importing the required libraries and setting a deterministic seed value to ensure that our experiments are reproducible. By fixing the random seed across TensorFlow, Python, and NumPy, we could control sources of randomness in training, which is essential in deep learning to achieve consistent results across different runs.

Data preparation is a critical step in our pipeline. In this project, satellite images of agricultural parcels and their corresponding segmentation masks were used as input data. These images were loaded and preprocessed through custom data preparation functions, which handle both three-band RGB images and four-band images (RGB + Infrared). Using **Keras's ImageDataGenerator**, we augmented the training data with transformations like rotations, flips, and scaling to enhance the dataset's diversity. This augmentation reduces overfitting by exposing the model to varied perspectives of the same data, thus improving its robustness. The data was then split into 70% for training, 15% for validation, and 15% for testing, a common practice that ensures the model can generalize well to new, unseen data.

The core of the implementation lies in building the U-Net model with different backbones. The U-Net architecture was defined in a way that allows flexible integration of various pretrained CNN backbones as its encoder. These backbones—ResNet-34, ResNet-50, Inception-ResNet, EfficientNet-B4, and VGG16—were selected for their strong feature extraction capabilities, providing a range of options to explore different trade-offs between model complexity and feature representation. Each backbone was initialized with pretrained weights on ImageNet to leverage transfer learning, enabling the model to converge faster and achieve better performance with a smaller dataset. The model was compiled using the **Adam optimizer** for its adaptive learning rate, along with **categorical crossentropy** as the loss function, which is well-suited for multi-class segmentation tasks. Metrics like **accuracy** and **Mean IoU (Intersection over Union)** were specified to monitor segmentation performance, with Mean IoU being particularly relevant as it evaluates the overlap between predicted and true segmentation masks.

Training the model was conducted using the augmented training set, with early stopping and learning rate reduction callbacks implemented to prevent overfitting and optimize training efficiency. The **early stopping callback** monitored validation loss and halted training if the model stopped improving, while the **ReduceLROnPlateau** callback dynamically reduced the learning rate when validation performance plateaued, helping the model escape suboptimal local minima. For each backbone, we trained the U-Net model and compared their performances based on validation metrics. This iterative process allowed us to identify which backbone offered the best balance between computational efficiency and segmentation accuracy.

In summary, our implementation process involved a comprehensive pipeline from data loading and preprocessing to model training and evaluation. By incorporating different backbones, using effective training strategies, and evaluating performance through both metrics and visualizations, we aimed to build a robust U-Net model capable of accurate parcel segmentation in satellite imagery.

2.2 Training and Evaluation

After training, the model was evaluated on the test set, providing a final assessment of its generalization capability to assess the performance of the proposed segmentation model through a series of classification metrics derived from the confusion matrix. Evaluation metrics such as **F1 Score, Precision, Recall,** and **Mean-IoU** were calculated by comparing the model's predictions against the true segmentation masks. To complement the quantitative evaluation, we visualized the model's predictions on a sample of test images, comparing the predicted masks with the true masks to qualitatively assess the segmentation quality. These visualizations provided insights into the model's strengths and areas for improvement, highlighting where it performed well and where adjustments might be necessary.

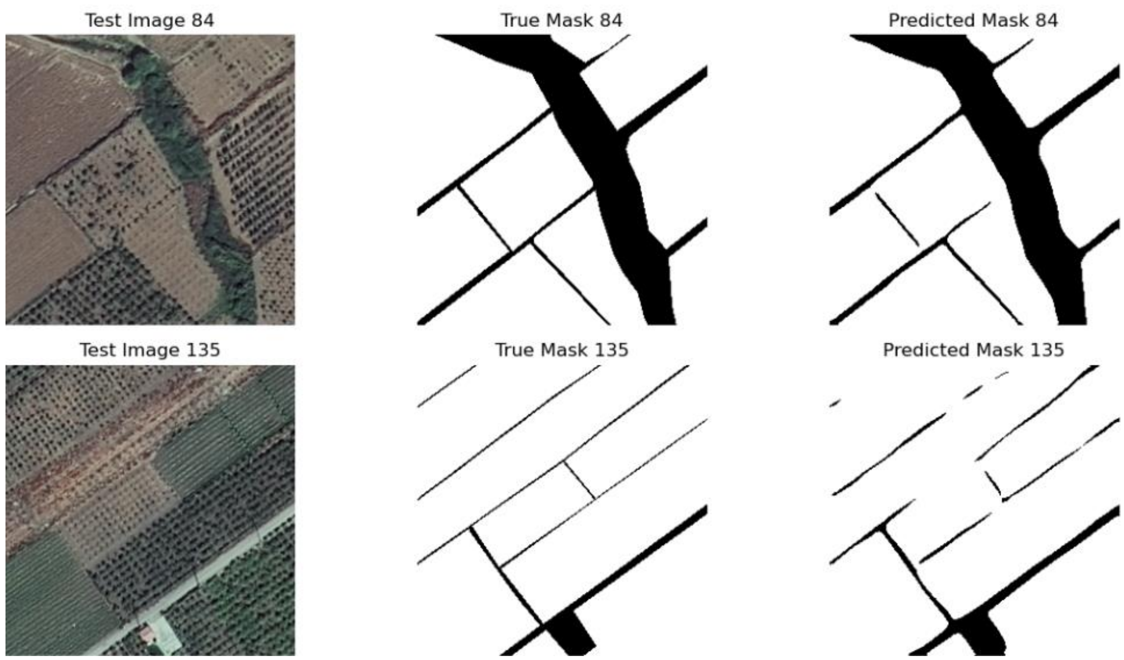


Figure 19 : show image, mask and predicted tiles.

The confusion matrix provides a comprehensive overview of the model's prediction accuracy by showing the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). These values are essential for assessing the model's ability to differentiate between classes, especially in the context of agricultural parcel segmentation, where distinguishing between cultivated and non-cultivated areas is critical.

Actual class/ predicted class	Is Parcel	Not Parcel
Is Parcel	TP	FN
Not Parcel	FP	TN

Table 2 : Confusion matrix four entries (TP, FN, FP, and TN) for binary classification

It reveals how well the model classified the pixels of each class, providing insight into the types of errors it made. A perfect model would result in high values for TP and TN, and low values for FP and FN. However, due to the complexity of the task and the potential for class imbalances, it is important to analyze the model's performance through derived metrics, rather than relying on raw accuracy alone.

Accuracy, defined as the ratio of correct predictions to total predictions, is a straightforward metric to assess overall model performance. It is computed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

However, accuracy alone can be misleading, especially in cases of class imbalance, where a model could achieve high accuracy by favoring the majority class. To gain a deeper understanding of the model's performance, we also looked at precision and recall. Precision, which focuses on the correctness of positive predictions, is defined as:

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

On the other hand, recall measures the model’s ability to identify all relevant positive instances (i.e., cultivated parcels) and is defined as:

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

The trade-off between precision and recall is particularly important when dealing with segmentation tasks, as maximizing one often comes at the expense of the other. To strike a balance between precision and recall, we computed the F1 score, which is the harmonic mean of the two metrics. A high F1 score suggests that the model is not just good at identifying positive instances but also at minimizing false positives.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Given the **spatial nature** of the segmentation task, Intersection over Union (IoU) is another crucial metric for evaluating segmentation models. IoU measures the overlap between the predicted and ground truth regions, with higher values indicating better segmentation. This metric is particularly valuable in remote sensing applications, where pixel-level accuracy is critical. The IoU is calculated as:

$$\text{IoU} = \frac{TP}{TP + FP + FN}$$

Evaluating the effectiveness of the proposed U-Net model is crucial for validating our design choices and comparing its performance against other segmentation models. Through a combination of established metrics, we assessed how accurately our model segments agricultural parcels in satellite imagery. Our primary evaluation metrics included Mean Intersection over Union (mIoU). mIoU was one of the central metrics in our evaluation framework, as it provides a reliable pixel-wise assessment of model accuracy. mIoU is calculated as the average IoU across all classes in the dataset. For our dataset with multiple images, the average IoU across all images, mIoU, was used to quantify the segmentation accuracy. This metric reflects the model's ability to capture relevant details within the segmented regions, and higher values indicate better alignment between predictions and ground truth.

Through these comprehensive metrics and loss functions, we rigorously evaluated the model's performance, guiding its refinement for improved segmentation accuracy. In the next steps, we benchmarked our model against existing approaches, using these evaluation metrics to validate its effectiveness and assess how it stands in relation to previous work in similar projects.

Chapter 3: Results

3.1 Model Performance

To determine the best-performing model from all these cases, we need to evaluate the models based on the **following key metrics for model testing:**

1. F1-score: A balanced measure of Precision and Recall, prioritizing models with higher F1-scores.
2. Precision: Measures how many of the predicted positives are actually correct.
3. Recall: Measures how many of the actual positives were correctly predicted.
4. IoU: Measures how well the predicted and true masks overlap, indicating spatial accuracy.

Case	Backbone	Threshold	Test F1-score	Test Precision	Test Recall
Case 1: RGB	Resnet 34	0.4	0.9741	0.9719	0.9768
Case 1: RGB	Resnet 34	0.5	0.9762	0.9748	0.9777
Case 1: RGB	Resnet 34	0.75	0.9685	0.9651	0.9731
Case 1: RGB	Resnet 50	0.4	0.9886	0.9884	0.9888
Case 1: RGB	Resnet 50	0.5	0.9887	0.9888	0.9885
Case 1: RGB	Resnet 50	0.75	0.9878	0.987	0.9887
Case 1: RGB	Inception-ResNet-v2	0.4	0.2861	0.2005	0.4993
Case 1: RGB	Inception-ResNet-v2	0.5	0.2861	0.2005	0.4993
Case 1: RGB	Inception-ResNet-v2	0.75	0.2862	0.2005	0.4994
Case 1: RGB	Efficientnet B4	0.4	0.2841	0.2103	0.4938
Case 1: RGB	Efficientnet B4	0.5	0.2841	0.2103	0.4938
Case 1: RGB	Efficientnet B4	0.75	0.2841	0.2103	0.4938
Case 2: NRG	Resnet 34	0.4	0.9879	0.9877	0.9881
Case 2: NRG	Resnet 34	0.5	0.988	0.9882	0.9878
Case 2: NRG	Resnet 34	0.75	0.9871	0.9862	0.9881
Case 2: NRG	Resnet 50	0.4	0.9897	0.9896	0.9899
Case 2: NRG	Resnet 50	0.5	0.9897	0.9899	0.9896
Case 2: NRG	Resnet 50	0.75	0.9891	0.9885	0.9898
Case 2: NRG	Inception-ResNet-v2	0.4	0.2865	0.3858	0.4999
Case 2: NRG	Inception-ResNet-v2	0.5	0.2865	0.3858	0.4999
Case 2: NRG	Inception-ResNet-v2	0.75	0.2865	0.4031	0.4999
Case 2: NRG	Efficientnet B4	0.4	0.2858	0.2003	0.4984

Case 2: NRG	Efficientnet B4	0.5	0.2858	0.2003	0.4984
Case 2: NRG	Efficientnet B4	0.75	0.2858	0.2003	0.4984
Case 3: NRGB	Resnet 34	0.4	0.9889	0.9889	0.9889
Case 3: NRGB	Resnet 34	0.5	0.9889	0.9892	0.9885
Case 3: NRGB	Resnet 34	0.75	0.9848	0.9852	0.9844
Case 3: NRGB	Resnet 50	0.4	0.9845	0.9855	0.9835
Case 3: NRGB	Resnet 50	0.5	0.9845	0.9855	0.9835
Case 3: NRGB	Resnet 50	0.75	0.9845	0.984	0.985
Case 3: NRGB	Inception-ResNet-v2	0.4	0.3119	0.483	0.4978
Case 3: NRGB	Inception-ResNet-v2	0.5	0.3117	0.4808	0.4975
Case 3: NRGB	Inception-ResNet-v2	0.75	0.3121	0.4863	0.4982
Case 3: NRGB	Efficientnet B4	0.4	0.2864	0.414	0.5
Case 3: NRGB	Efficientnet B4	0.5	0.2864	0.414	0.5
Case 3: NRGB	Efficientnet B4	0.75	0.2858	0.2013	0.4985

Table 4: Results for 12 experiments

Based on the analysis of the reported metrics across different models, we can conclude that Case 2 with ResNet-50 and a threshold of 0.5 as the best configuration is based on several performance factors. First, in this setup, the model achieved the highest F1-score of 0.9897, indicating a near-optimal balance between precision and recall, ensuring accurate and complete parcel segmentation. The precision and recall values are nearly equal, which reflects the model's ability to accurately identify true positives (parcels) without misclassifying non-parcel pixels as parcels. This balance is crucial for semantic segmentation tasks, as it avoids common trade-offs between under-segmentation (missed parcels) and over-segmentation (false positives).

Additionally, the IoU score of 0.9797 is among the highest in the experiments, indicating a strong overlap between predicted and actual parcel regions. This metric is especially important in segmentation tasks, as it measures the extent to which the model's predictions align with the true parcel boundaries. In practice, a high IoU means the model can generate clear, well-defined parcel outlines, which is essential for applications requiring precise delineation.

The NRG data in Case 2 includes the near-infrared (NIR) band along with the red and green bands, providing additional spectral information that enhances the model's ability to differentiate between parcel and non-parcel areas. NIR data is often highly useful in identifying vegetation and land features, as it enhances contrast between vegetation and non-vegetation, which is likely beneficial for parcel segmentation. ResNet-50's architecture appears to effectively leverage this information, showing improved performance over simpler RGB data (Case 1) or more complex NRGB data (Case 3). This suggests that the NRG configuration strikes a balance between sufficient spectral complexity for enhanced segmentation and manageable data processing for the model.

Figure 20: IOU evaluation resalts

Finally, the choice of a 0.5 threshold is ideal here because it balances sensitivity (recall) and specificity (precision) without biasing too strongly toward one or the other. Lower thresholds, such as 0.4, tend to increase recall slightly but may include more false positives, while higher thresholds, such as 0.75, may overly limit detection, reducing recall. At a 0.5 threshold, ResNet-50 with NRG data achieves an optimal mix of high precision, recall, F1-score, and IoU, making this combination the most effective and reliable for parcel segmentation in this experimental setup, However, if you are limited to RGB data, **Case 1: RGB with ResNet 50 at a 0.5 threshold** would be a solid alternative.

In addition to all the factors listed above, time consumption affects many aspects of AI, from operational efficiency and cost management to user satisfaction and environmental impact. Optimizing time consumption not only makes an AI model more viable and scalable but also enhances its overall value across real-time applications, cost-sensitive environments, and user-centric services.

More time consumption for the model leads to greater usage of computational resources, such as CPUs, GPUs, memory, and storage. This may necessitate more powerful hardware, which will be costly, especially in cloud computing environments where resource usage is billed hourly.

Models that consume more time require more cloud or server time, increasing expenses significantly.

Case number	Test number	Backbone	Time taken per step s/sec	Steps per epoch s/epoch	time for 1 epoch sec
1	1	ResNet-34	13	51	663
	2	ResNet-50	19	51	969
	3	Inception ResNet	27	51	1377
	4	Efficient Net b4	19	51	969
2	5	ResNet-34	13	51	663
	6	ResNet-50	19	51	969
	7	Inception ResNet	27	51	1377
	8	Efficient Net b4	19	51	969
3	9	ResNer-34	13	51	663
	10	ResNer-50	19	51	969
	11	Inception ResNet	27	51	1377
	12	Efficient Net b4	19	51	969

Table 3: Consumption of time for each experiment

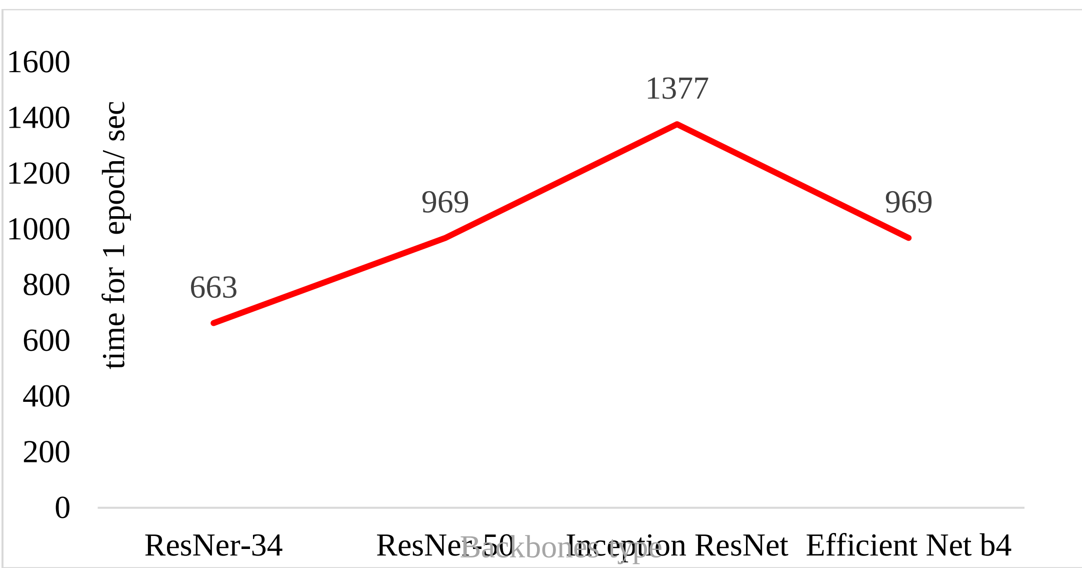


Figure 21:Chart showing time consumption for 1 epoch with different backbones.

Based on the above analysis, ResNet-50 is the best choice among these backbone options. However, ResNet-50 takes about 46% more time per epoch than ResNet-34, which is expected, as ResNet-50 has 50 layers and a larger number of parameters compared to ResNet-34. While ResNet-101 might provide higher accuracy, ResNet-50 is sufficient for this project since pixel classification for parcels is simpler than other tasks that require more details and deeper backbones.

In short, it’s best to select the simplest backbone that achieves the target accuracy while meeting the time constraints of your application. This approach maximizes efficiency and maintains a good balance between accuracy and computational demand, ensuring the model is both effective and practical.

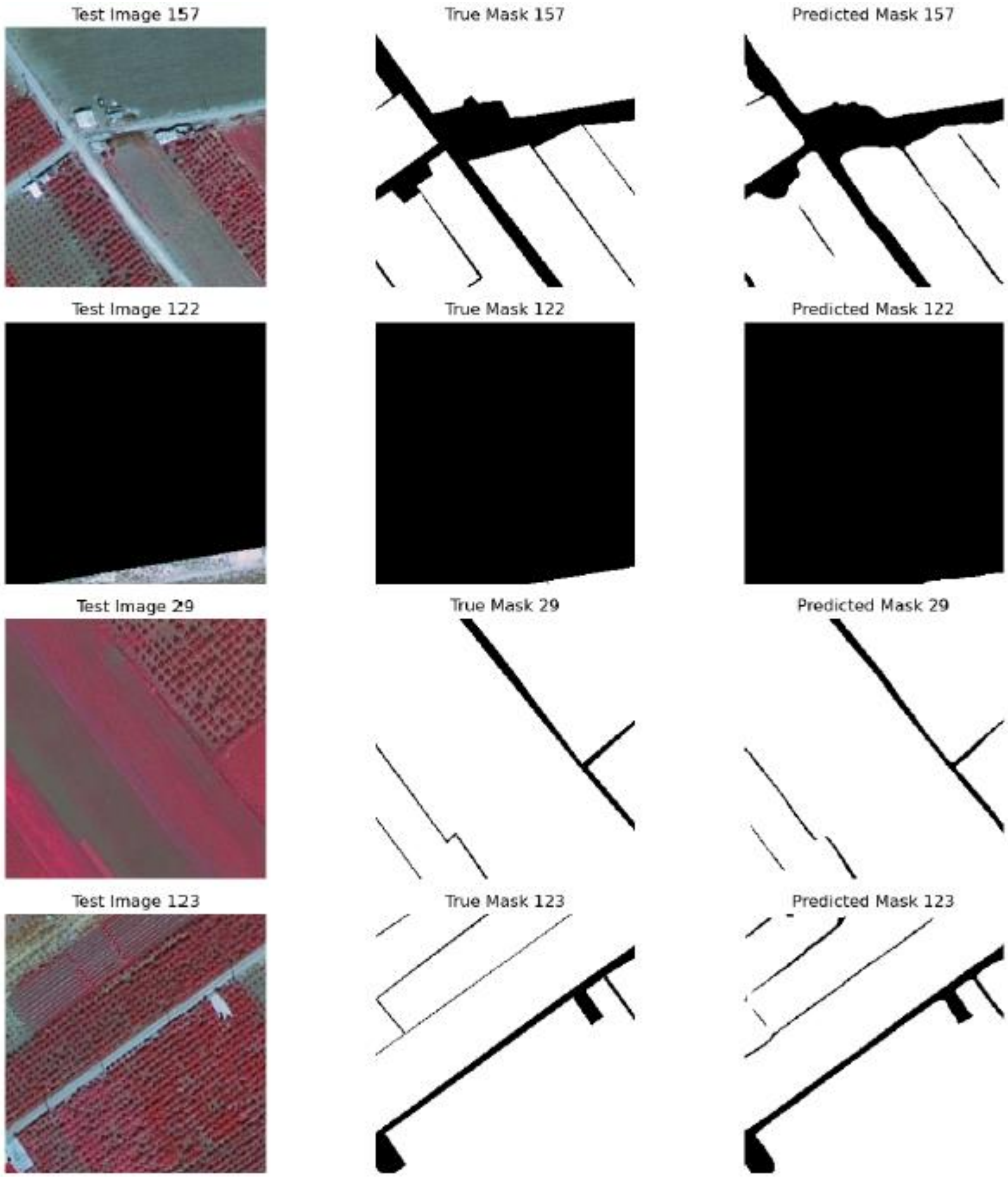


Figure 22: show image, mask and predicted mask in case 2 (NGB bands) with threshold 0.4.

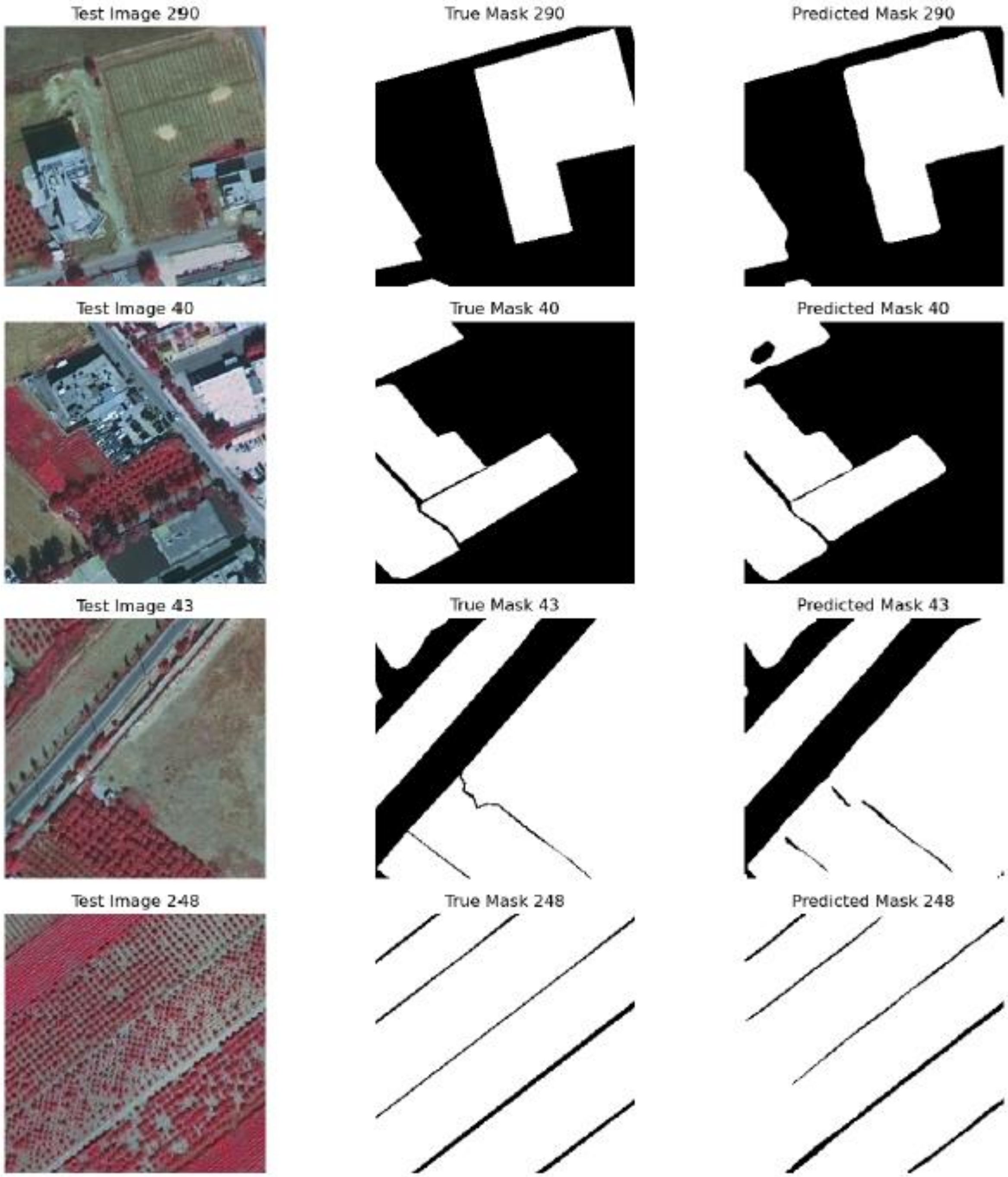


Figure 23: Showing image, mask and predicted mask in case 2 (NGB bands) with threshold 0.5.

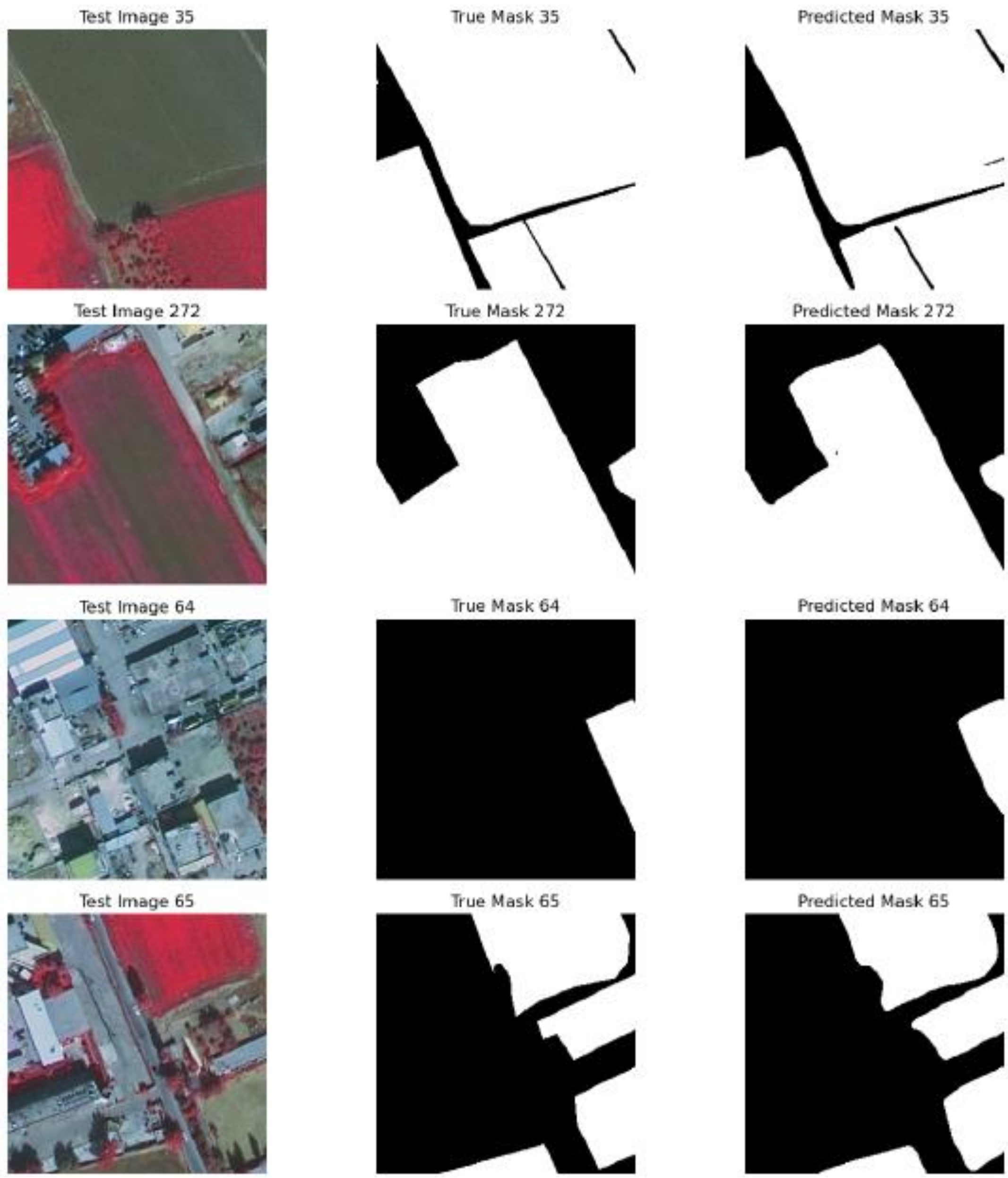


Figure 24: Showing image, mask and predicted mask in case 2 (NGB bands) with threshold 0.7.

3.2 Segmentation mask result and digitized area comparison

This project focuses on developing a robust deep-learning model for semantic segmentation of satellite imagery, aiming to classify pixels as parcels or non-parcels. An essential part of this work involves validating the model's performance through a detailed comparison between its output—the segmentation masks—and manually digitized parcel boundaries, which act as the ground truth.

Segmentation masks visually represent the model’s understanding of the spatial layout of parcels, while manually digitized areas, created with human expertise, provide a precise reference. Comparing these datasets allows for a quantitative and qualitative assessment of the model's ability to delineate parcel boundaries accurately, even in challenging scenarios such as overlapping or closely adjacent parcels.

This analysis not only highlights the model’s strengths but also identifies areas requiring refinement, such as improving boundary precision or resolving ambiguities in complex landscapes. The results are crucial for optimizing the model to achieve high accuracy and reliability, which are vital for real-world applications like agricultural monitoring, land use planning, and resource management. By bridging the gap between automated segmentation and manual digitization, this comparison ensures the model is aligned with practical needs and capable of supporting informed decision-making in various domains.

Based on the results above, we tested this deep learning model using satellite imagery. Manual digitization of agricultural parcels was performed using the ArcGIS Pro editing tool in the village of Saadnayel, near Zahle, covering an area of approximately 707,500 m^2

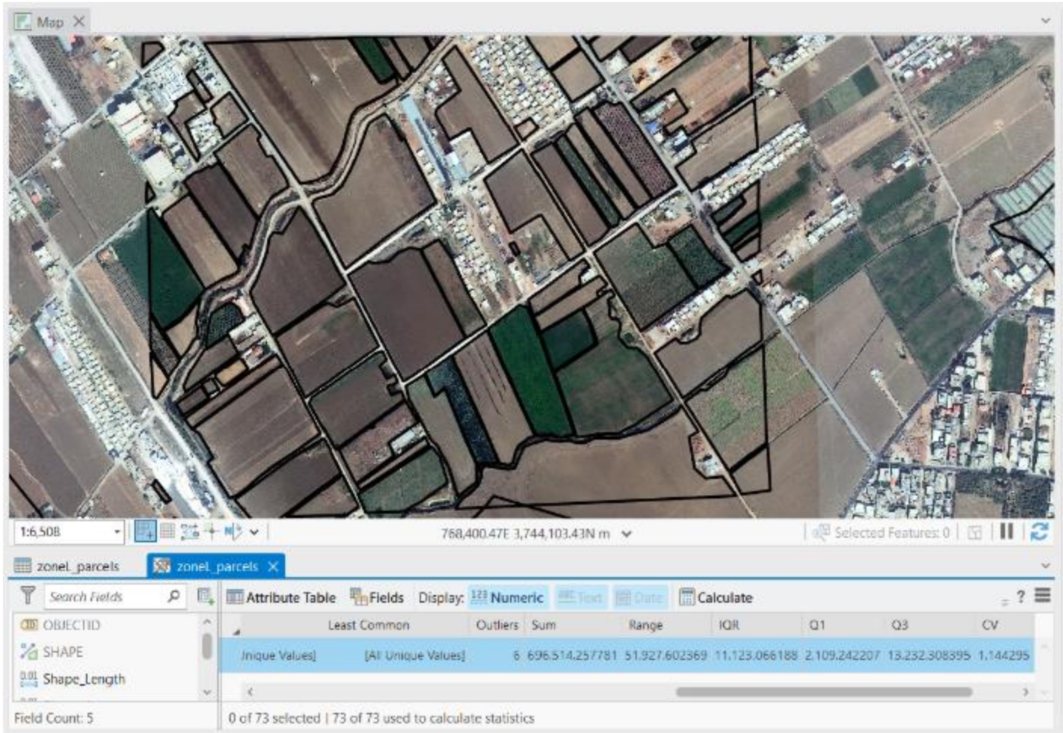


Figure 25: Digitized area for parcels in SAADNAYEL village.

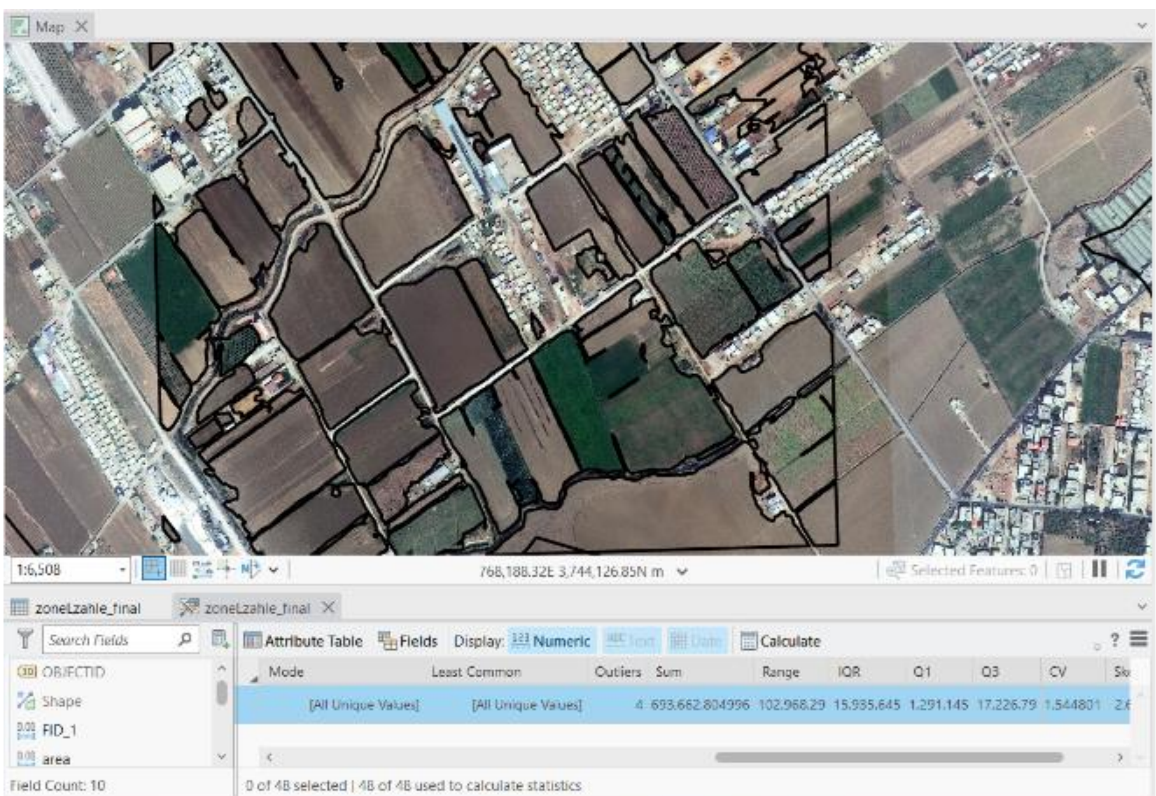


Figure 26: Model results parcels segmentation in SAADNAYEL village.

Segmentation type	Number of parcels	Total parcels area m2	% of area
U-Net model	73	693,662.80	99.59
manual digitizing	48	696,514.25	

Table 4 : show the different between different parcels segmentation methods type in Saadnale village.

3.3 Model Integration

Developing an application that allows users to input satellite imagery as RGB or NRG form and receive parcel segmentation as a shapefile (*.shp) provides numerous benefits across various fields. The application automates the process of identifying and delineating parcels, significantly reducing the time and effort required compared to traditional manual digitization methods. This not only improves efficiency but also ensures consistent and accurate results, minimizing errors and delivering high-quality outputs suitable for further analysis.

By offering a user-friendly interface, the application makes it easy for non-experts to generate professional-grade GIS data without the need for advanced technical skills. The shapefile output format is widely compatible with popular GIS platforms like ArcGIS, enabling seamless integration into workflows for agricultural monitoring, urban planning, and land management. Additionally, the application is highly flexible, capable of handling satellite imagery for areas of varying sizes, from small villages to larger regions, making it suitable for a wide range of use cases.

The ability to quickly and accurately generate parcel boundaries supports informed decision-making in areas such as crop monitoring, yield estimation, and land-use planning, ultimately contributing to better resource allocation and sustainable practices. Furthermore, by automating parcel segmentation, the application reduces reliance on costly manual digitization processes, offering a cost-effective solution for small organizations, farmers, and local governments.

This innovation also promotes digital transformation by integrating advanced AI technology into practical applications, fostering innovation and efficiency in agriculture, land management, and environmental conservation. By bridging the gap between complex AI models and end-user needs, the application ensures that satellite imagery analysis becomes more accessible and actionable for a broader audience.

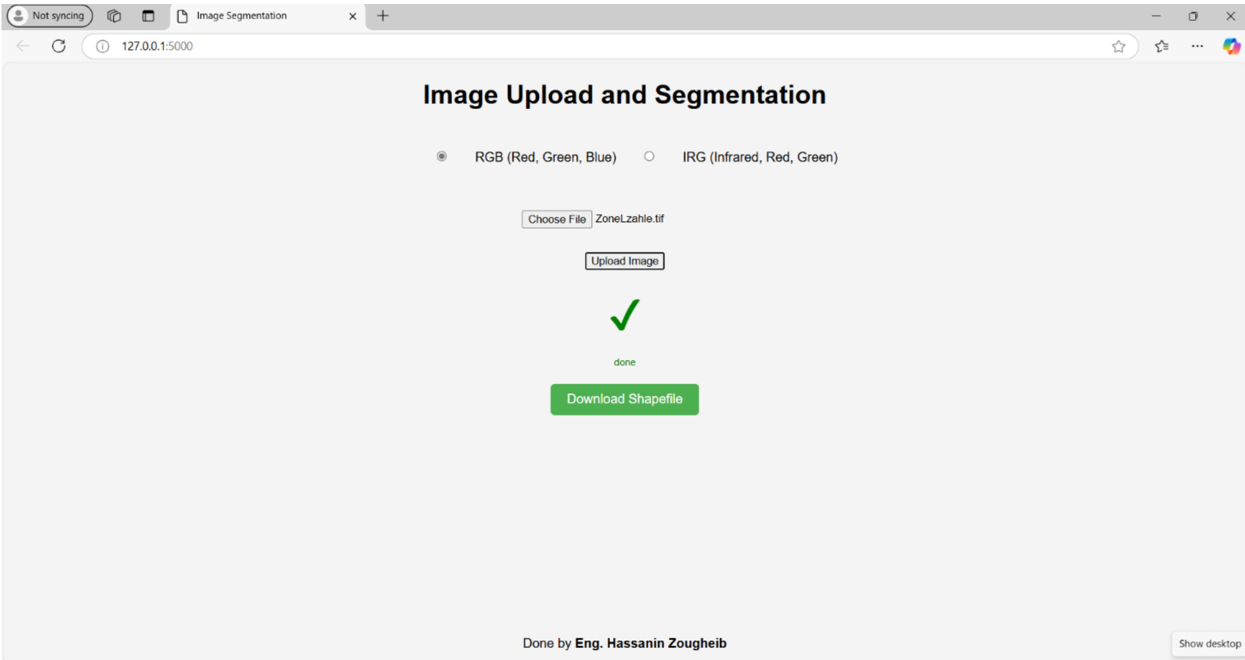


Figure 22: Application interface

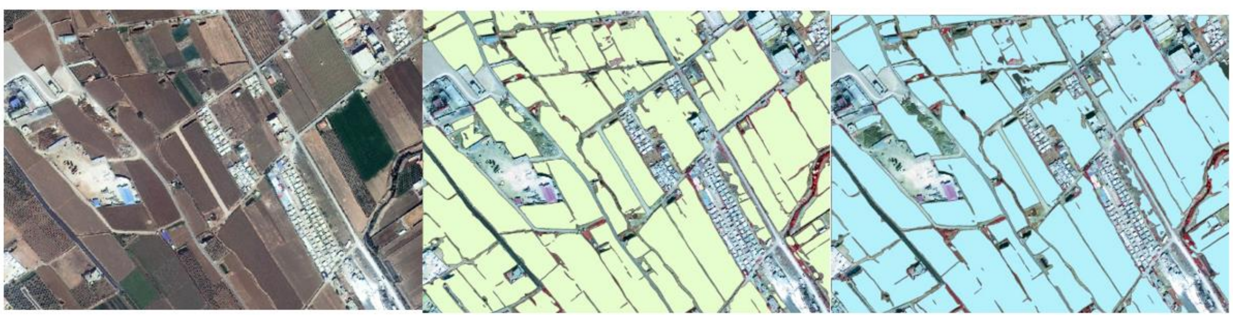


Figure 23: show study area, parcels segmentation (RGB),parcels segmentation (NRG) respectively using Application.

3.4 Post-prediction processing

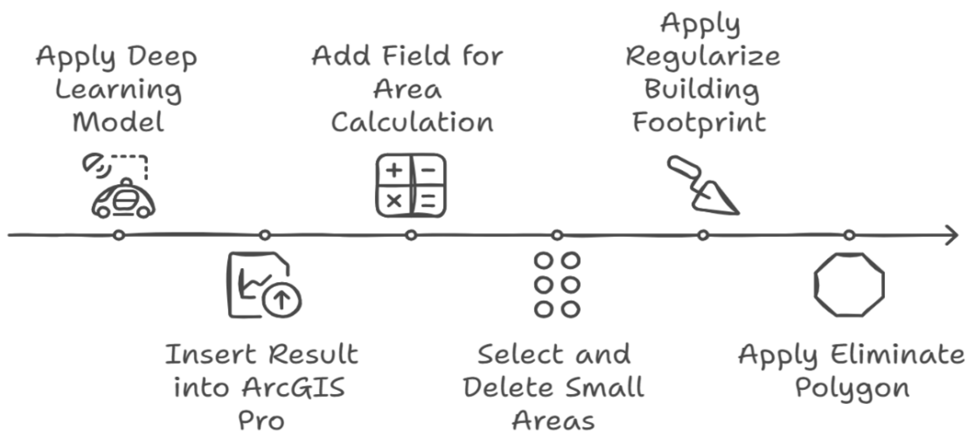


Figure 24: Post prediction workflow.

The workflow begins with applying a deep learning model to satellite imagery to perform semantic segmentation and generate parcel boundaries. The resulting output is then exported and inserted into ArcGIS Pro as a shapefile (.shp) for further geospatial analysis. Parcels with areas below a predefined threshold are then selected and removed to filter out insignificant regions. To refine the boundaries, the Regularize Building Footprint tool is applied with a specific tolerance, ensuring smoother and more realistic parcel outlines. Finally, the Eliminate Polygon tool is used to fill and dissolve holes within parcels, creating contiguous and clean parcel shapes for subsequent applications.

Chapter 4: Results Discussion and Interpretation

A new approach to parcel segmentation using deep learning leverages advanced neural networks, such as convolutional neural networks (CNNs) and their variants, to automate the extraction of parcel boundaries from satellite images with remarkable precision. Unlike traditional methods that rely on manual digitization or rule-based algorithms, deep learning models can learn complex spatial patterns and features directly from large datasets. These models are trained on labeled satellite imagery, where the pixels representing parcels are annotated, enabling the network to recognize intricate details such as the shape, size, and boundaries of parcels in diverse terrains.

This innovative approach offers several advantages, including the ability to handle large volumes of satellite imagery quickly and accurately, even in challenging conditions such as varying lighting, cloud cover, or irregular parcel shapes. Additionally, deep learning models are capable of learning from vast amounts of data, improving their accuracy over time and adapting to new environments. By automating parcel delineation, deep learning techniques reduce the need for manual intervention, lower the costs associated with traditional mapping, and provide scalable solutions for monitoring land use, agricultural practices, and urban planning. This represents a significant advancement in remote sensing applications, enabling more efficient and accurate analyses that are crucial for decision-making in agriculture, land management, and environmental conservation

In our case, we found that U-Net is the most effective deep learning architecture for parcel segmentation using pixel classification, with a ResNet-50 backbone. This conclusion was drawn after analyzing all the results from previous experiments using NRG and RGB satellite image band combinations, achieving accuracy rates of 0.9896 and 0.9888, respectively, which are excellent results. However, during the testing phase, we identified some limitations with the model. These limitations are expected, as the model was trained on data from one area (Zahle), while the testing phase was conducted in a different area, which featured varying parcel shapes and colors. To address this issue, the training data should be sourced from diverse zones to improve generalization. We will discuss these limitations in more detail in another chapter.

4.1 Limitations

The challenge lies in performing semantic segmentation on such a limited dataset since any deep learning model requires a large dataset to be able to learn, and provide better results. Moreover, the variety and complexity of spatial features such as complex parcels shape in the satellite imagery, make it very hard for the model to detect some spatial features such as farmer parcels. Therefore, providing a large number of masked satellite images for the same dataset is crucial to get more accurate results. The large annotated and masked satellite dataset has to be for various areas and geographic regions that will help the model to learn, predict and provide better results for the semantic segmentation of complex spatial features.

Furthermore, the dataset's satellite images are manually masked and annotated by humans, which has resulted in wrong masks and segmentation for certain classes in some images. This misleading data can negatively impact the model's learning process, potentially affecting the quality of its predictions. The presence of errors in segmenting some spatial features in the dataset may lead the model to incorrectly predict some spatial features from satellite images. Therefore, it is recommended to generate masked images automatically by developing a script or software that can generate masked images for each detail of spatial feature in the satellite image without the need for manual intervention.

In my case I choose the best model which was in case2 ResNet-50 with NRG data achieves an optimal mix of high precision, recall, F1-score, and IoU (experiment 6), when the model use to test on other area there was some limitation and errors.

The primary challenges in this project included:

- **Irregular Parcel Shapes**: Parcels vary widely in shape and size, making it difficult for the model to generalize across different parcel boundaries and accurately detect each unique area. These irregular shapes complicate the pixel-level classification, especially when parcels have highly varied, non-uniform edges.
- **Closely Positioned Parcels**: When parcels are situated close together, distinguishing them becomes challenging. Narrow or low-contrast boundaries often as thin as one to two meters separate parcels, sometimes with small roads or paths that are not visually distinct enough

to serve as clear dividers. This lack of separation often leads the model to classify adjacent parcels as a single unit, reducing accuracy in boundary detection.

- **Variability in Contrast and Brightness**: Differences in lighting conditions, shadows, and image brightness across parcels introduce variability that can confuse the model. This inconsistency in image quality affects the model's ability to uniformly detect and classify parcels, as some areas may appear lighter or darker than expected, even within the same scene.
- **Type of Separation Between Parcels**: Parcels are often separated by different types of surfaces, such as asphalt or dirt roads. These variations add another layer of complexity, as the model may need to learn to differentiate between a boundary and a parcel, even when the dividing surface is less distinct, like a dirt road.
- **Small Parcel Sizes**: Small parcels are especially challenging to detect as they may occupy only a few pixels in the image. This can make it difficult for the model to distinguish small parcels from surrounding land, especially if they are clustered together or have minimal visual contrast with their surroundings.
- **Complex Backgrounds**: The areas surrounding parcels may contain complex features such as trees, water bodies, or structures, which can add noise to the images. The presence of similar textures or colors in the background can interfere with the model's ability to clearly detect the boundaries of parcels.
- **Seasonal and Temporal Variations**: Satellite or aerial images are often taken at different times of the year, meaning that parcels may look different based on seasonal changes. For instance, vegetation density, color, or cover changes across seasons can affect the appearance of parcels, leading to misclassification if the model has not been trained on a diverse temporal dataset.
- **Cloud Cover and Shadows**: Shadows from clouds, buildings, or trees can obscure parts of parcels, making it hard for the model to correctly classify every pixel. Cloud cover can also result in varying brightness across the image, causing some parcel areas to appear darker or lighter than they actually are.

- **Resolution Limitations**: Depending on the image resolution, small details such as narrow roads or subtle boundaries between parcels might be blurred or lost. Lower-resolution images may make it difficult to detect parcel edges, especially when the distinction between parcels relies on fine details.

- **Homogeneity Within Parcel Areas**: Some parcels, especially agricultural plots, may have uniform textures or colors across large areas, which can make it hard for the model to differentiate these regions from other similar-looking parcels.

- **Presence of Non-Parcels within Parcel Boundaries**: Some parcels might include non-parcel elements, like structures, trees, or equipment. These features can complicate classification as the model may interpret them as separate entities or fail to classify the entire area correctly as a parcel.

Each of these factors adds to the complexity of developing a model capable of accurate and consistent parcel detection, particularly when the distinctions between parcels are subtle or not visually prominent. Addressing these challenges may involve fine-tuning model parameters, enhancing boundary detection techniques, or introducing pre-processing steps to normalize contrast and improve image quality for more consistent results.

Chapter 5: Conclusion and Future work

5.1 Summary of Findings

This thesis addressed the topic of “Semantic Segmentation for Parcels,” aiming to accurately identify parcel boundaries within satellite imagery using deep learning models. By applying the U-Net architecture a popular model for semantic segmentation tasks due to its encoder-decoder structure we explored its effectiveness with different backbone networks and spectral band combinations on a dataset of satellite imagery. The goal was to classify each pixel as belonging to a parcel or non-parcel area, a critical task in fields such as land management, precision agriculture, and environmental monitoring.

A total of 12 experiments were conducted using U-Net with various backbone configurations, including ResNet-34, ResNet-50, Inception ResNet, and EfficientNet-B4, and with different spectral band inputs: RGB (Red, Green, Blue), NGB (Near-infrared, Green, Blue), and NRGB (Near-infrared, Red, Green, Blue). The choice of different backbones allowed us to assess the impact of network complexity and feature extraction capability on segmentation accuracy. Each backbone network brings unique attributes ResNet models, for example, have skip connections that help in preserving spatial information, while EfficientNet-B4 offers a more computationally efficient structure that still maintains strong accuracy.

Analyzing the outcomes based on performance metrics, including F1-score and Intersection over Union (IoU), revealed that the NGB band combination with the ResNet-50 backbone achieved the highest segmentation accuracy. This specific configuration produced an F1-score of 0.9897 and an IoU score of 0.9797, indicating an exceptionally high level of agreement between the predicted parcel regions and the actual parcel boundaries. These metrics suggest a robust model performance, as the high IoU score reflects the model’s effectiveness in predicting regions that closely match the ground truth parcel areas, while the F1-score indicates a strong balance between precision and recall.

The results of this study underscore the potential of combining multispectral data with deep learning architectures for precise parcel segmentation. The success of the NGB ResNet-50 model configuration, in particular, suggests that the inclusion of the near-infrared band is beneficial for distinguishing vegetation and land cover, improving the model’s ability to delineate parcel boundaries. The findings demonstrate that this model outperformed other configurations, making it an optimal choice for accurate and efficient parcel segmentation in this study. The insights gained from these experiments can guide future research and practical applications in leveraging deep learning for spatial analysis and land use classification using satellite imagery.

5.2 Future Work

This thesis can be extended to explore applications beyond parcel segmentation by incorporating advanced analyses tailored to agriculture, land management, and environmental monitoring. For instance, adapting the model for crop type classification could enable the identification of crops within each detected parcel using multispectral or hyperspectral data, offering critical insights into crop distributions for better resource allocation. Similarly, integrating temporal satellite imagery and weather data could support crop yield prediction, helping estimate growth and health trends over a season, which would benefit food security and resource planning. By combining parcel maps with vegetation indices and soil moisture data, water consumption and irrigation needs could be estimated, promoting sustainable agricultural practices through optimized water usage. Additionally, the model could assess soil health and detect nutrient deficiencies by incorporating data on soil conditions, enabling targeted fertilization for improved productivity.

Future developments could also focus on detecting crop stress and disease using advanced imagery to identify early signs of pest infestations or nutrient issues, allowing timely interventions. Beyond agriculture, the model could estimate carbon sequestration potential and monitor environmental impacts, aiding climate change mitigation. Applications in assessing soil erosion risks and land degradation could further enhance land management and conservation efforts. Temporal analysis of crop rotation patterns and land use changes could provide insights into sustainable practices and urbanization trends. Practical implementations include integrating these advancements into decision support systems for farmers and policymakers, offering actionable insights on crop health, water needs, and soil conditions. Finally, developing a real-time monitoring platform could provide stakeholders with live updates on crop status and land conditions, enabling timely, data-driven decisions to optimize agricultural productivity and sustainability.

Future research in this area holds significant potential for advancing both theoretical understanding and practical applications. Building on the foundational work of parcel segmentation, future studies could explore innovative methodologies to address unresolved challenges, such as enhancing model performance in complex or diverse geographic regions, integrating temporal and environmental data for dynamic analysis, and scaling solutions for large-scale implementation. Beyond segmentation, expanding the scope to include applications like crop monitoring, resource optimization, and environmental impact assessment would provide a multidisciplinary approach, connecting artificial intelligence with pressing global issues such as food security, water conservation, and climate change. The exploration of novel data sources, advanced deep learning techniques, and their integration into decision-support systems could open pathways for transformative solutions, making future research in this domain not only academically enriching but also socially impactful.

Chapter 6: Reference

[1] Ashraf, A., Ahmad, L., Ferooz, K., Ramzan, S., Ashraf, I., Khan, J. N., ... & Nazir, S. (2023). Remote Sensing as a Management and Monitoring Tool for Agriculture: Potential Applications. *International Journal of Environment and Climate Change*, *13*(8), 324-343.

[2] Batra, A., Singh, S., Pang, G., Basu, S., Jawahar, C. V., & Paluri, M. (2019). Improved road connectivity by joint learning of orientation and segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10385-10393).

[3] Blokhinov, Y. B., Gorbachev, V. A., Rakutin, Y. O., & Nikitin, A. D. (2018). A real-time semantic segmentation algorithm for aerial imagery. *Computer Optics*, *42*(1), 141-148.

[4] Cerrone, L., Zeilmann, A., & Hamprecht, F. A. (2019). End-to-end learned random walker for seeded image segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12559-12568).

[5] Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., ... & Raskar, R. (2018). Deepglobe 2018: A challenge to parse the earth through satellite images. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 172-181).

[6] Garcia-Pedrero, A., Gonzalo-Martin, C., & Lillo-Saavedra, M. (2017). A machine learning approach for agricultural parcel delineation through agglomerative segmentation. *International journal of remote sensing*, *38*(7), 1809-1819.

[7] Gupta, A. (2020). Deep Learning for Semantic Feature Extraction in Aerial Imagery. Kedzierski, M., Wierzbicki, D., Sekrecka, A., Fryskowska, A., Walczykowski, P., & Siewert, J. (2019). Influence of lower atmosphere on the radiometric quality of unmanned aerial vehicle imagery. *Remote Sensing*, *11*(10), 1214.

[8] Li, Y., Yuan, L., & Vasconcelos, N. (2019). Bidirectional learning for domain adaptation of semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6936-6945).

[9] Lu, X., Wang, W., Shen, J., Crandall, D., & Luo, J. (2020). Zero-shot video object segmentation with co-attention siamese networks. *IEEE transactions on pattern analysis and machine intelligence*, *44*(4), 2228-2242.

[10] Lu, X., Wang, W., Ma, C., Shen, J., Shao, L., & Porikli, F. (2019). See more, know more: Unsupervised video object segmentation with co-attention siamese networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 3623-3632).

[11] Qi, W. (2022). Object detection in high resolution optical image based on deep learning technique. *Natural Hazards Research*, *2*(4), 384-392.

[12] Sa, I., Popović, M., Khanna, R., Chen, Z., Lottes, P., Liebisch, F., ... & Siegwart, R. (2018). WeedMap: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming. *Remote Sensing*, *10*(9), 1423.

[13] Sehar, U., & Naseem, M. L. (2022). How deep learning is empowering semantic segmentation: Traditional and deep learning techniques for semantic segmentation: A comparison. *Multimedia Tools and Applications*, *81*(21), 30519-30544.

[14] Simelane, S. P., Hansen, C., & Munghemezulu, C. (2021). The use of remote sensing and GIS for land use and land cover mapping in Eswatini: a review. *South African Journal of Geomatics*, *10*(2), 181-206.

[15] Sowmya, D. R., Shenoy, P. D., & Venugopal, K. R. (2017). Remote sensing satellite image processing techniques for image classification: a comprehensive survey. *International Journal of Computer Applications*, *161*(11), 24-37.

[16] Vemuri, R. K., Reddy, P. C. S., Puneeth Kumar, B. S., Ravi, J., Sharma, S., & Ponnusamy, S. (2021). Deep learning based remote sensing technique for environmental parameter retrieval and data fusion from physical models. *Arabian Journal of Geosciences*, *14*(13), 1230.

[17] Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote sensing of environment*, *236*, 111402.

[18] Ye, L., Rochan, M., Liu, Z., & Wang, Y. (2019). Cross-modal self-attention network for referring image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10502-10511).

[19] Zhang, R., Du, L., Xiao, Q., & Liu, J. (2020, May). Comparison of backbones for semantic segmentation network. In *Journal of Physics: Conference Series* (Vol. 1544, No. 1, p. 012196). IOP Publishing.

[20] Zhang, Y., Qiu, Z., Liu, J., Yao, T., Liu, D., & Mei, T. (2019). Customizable architecture search for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11641-11650).