

Master of Science Program
Photogrammetry and Geoinformatics
Master Thesis
Winter Term 2024/2025

Characterizing Corn Growth Stages from UAV Imagery in Guatemala

by

B. Eng. Emerson Javier Martinez García

Supervisors:

Prof. Dr. Sven Schneider

M. Sc. Carlos Navarro

Characterizing corn growth stages from UAV imagery in Guatemala

by

B. Eng. Emerson Javier Martínez García

A dissertation presented in partial fulfillment of the requirements for the degree of Master of Science in the Department of Geomatics, Computer Science and Mathematics, Stuttgart University of Applied Sciences

Declaration

The following Master thesis was prepared in my own words without any additional help. All used sources of literature are listed at the end of the thesis.

I hereby grant to Stuttgart University of Applied Sciences permission to reproduce and to distribute publicly paper and electronic copies of this document in whole and in part.

Stuttgart, 28.02.2025

(Emerson Javier Martínez García)

Approved by:

(Prof. Dr. Sven Schneider)

Acknowledgements

I would like to express my sincere gratitude to **Prof. Dr. Sven Schneider** from **HFT Stuttgart** for his invaluable guidance and mentorship throughout this research. I am also deeply grateful to **MSc. Carlos Navarro** and **MSc. Mario Chávez** from **Alliance Bioversity & CIAT** for continuous support and insightful feedback. A special thanks to all the **staff members of Alliance Bioversity & CIAT**, whose expertise and collaboration played a crucial role in shaping this work.

I extend my heartfelt appreciation to my **master colleagues, Monica Acosta, Johanna Esguerra, Arsalan Mukhtar, Leela Pathipati, Ashifur Rahaman, Meera Prajapati, Fatemeh Rafiei, and Jeeva Jeyaram**, for their unwavering support, motivation, and companionship throughout this journey. Their encouragement and friendship have been instrumental in overcoming challenges and celebrating achievements.

I also want to express my deepest gratitude to my **friends from my homeland, Nicaragua—Leandrus Torres, Manuel Sequeira, Sergio Bustamante, and Aarón Méndez**—who have always been there for me, offering their support, encouragement, and motivation despite the distance.

Above all, I dedicate this accomplishment to my **parents, Loida García and Emerson Martínez**, whose unconditional love and sacrifices have been my greatest source of strength. I also dedicate this to my **sisters, Loida Martínez, Ericka Martínez, and Xochilt Martínez** and my **uncle Zamir García**, for their endless encouragement and support. Lastly, I dedicate this to my **nephews and the one who is coming soon**, whose presence fills my heart with joy and inspiration. Their love and belief in me have been my driving force every single day.

Master Course Photogrammetry and Geoinformatics

Characterizing corn growth stages from UAV imagery in Guatemala

Abstract

Smallholder farmers struggle to monitor crop growth due to limited access to advanced agricultural technologies. This study enhances crop monitoring using UAV imagery and U-Net-based image segmentation to map fields and assess vegetation health. The model was applied in two contrasting environments, Totonicapán (highland) and Zacapa (lowland), achieving higher segmentation accuracy in highlands, where spectral contrast improves classification, while lowland segmentation faced challenges due to overlapping vegetation. Despite this, the model effectively extracted NDVI values for statistical analysis. The Shapiro-Wilk test confirmed normality in highland NDVI, while lowland NDVI showed minor deviations. A t-test ($t = 4.4823$, $p < 0.001$) confirmed statistically significant NDVI differences between regions, reflecting distinct agronomic conditions. Interestingly, lowlands exhibited higher yield (2 t/ha) than highlands (0.5 t/ha), despite lower NDVI values, suggesting NDVI alone is not a yield predictor. These findings highlight U-Net's potential for precision agriculture in supporting data-driven decision-making for smallholder farmers.

Keywords: U-Net, UAV Imagery, NDVI, T-Test, Crop Monitoring, Smallholder Farmers, Image Segmentation, Lowlands, Highlands, Remote Sensing.

Contents

DECLARATION	1
ACKNOWLEDGEMENTS	2
ABSTRACT	3
CONTENTS	4
LIST OF FIGURES	6
LIST OF TABLES	7
ABBREVIATIONS	8
1. INTRODUCTION.....	9
2. OBJECTIVES	11
2.1 GENERAL.....	11
2.2 SPECIFICS	11
3. SCIENTIFIC OVERVIEW.....	12
3.1 REMOTE SENSING IN AGRICULTURE	12
3.2 MACHINE LEARNING IN CROP ANALYSIS	16
3.3 CHARACTERIZATION OF GROWTH STAGES USING NDVI.....	26
3.4 TEMPORAL AND SPATIAL ANALYSIS OF NDVI IN PHENOLOGICAL CHANGES	30
3.5 REGIONAL CONTEXT: ZACAPA AND TOTONICAPÁN	32
4. METHODOLOGY.....	37
4.1 STUDY AREA	37
4.2 DATA ACQUISITION	38
4.3 IMAGE PREPROCESSING.....	42
4.4 MACHINE LEARNING MODEL	44
4.5 NDVI ANALYSIS	48
4.6 AGRONOMIC DATA	50
5. RESULTS.....	51
5.1 PERFORMANCE OF THE U-NET MODEL, IMAGE SEGMENTATION.....	51
5.1.1 Training	51
5.1.2 Validation.....	54
5.2 REGIONAL COMPARISON AND T-TEST RESULTS	58
5.2.1 NDVI Distribution for Two Locations:	58

5.2.2	<i>Statistical Test Results</i>	59
5.3	NDVI METRICS ANALYSIS	61
5.3.1	<i>Mean NDVI Trends</i>	61
5.3.2	<i>Minimum and Maximum NDVI</i>	61
5.3.3	<i>Dispersion and Variability</i>	62
5.4	AGRONOMIC DATA INTEGRATION	63
6.	CONCLUSIONS	64
7.	REFERENCES.....	65

List of Figures

FIGURE 1 NDVI ACROSS CROP GROWTH STAGES	15
FIGURE 2 CROP VARIABILITY USING NDVI FROM HIGH-RESOLUTION UAV IMAGE	16
FIGURE 3 U-NET ARCHITECTURE DIAGRAM	20
FIGURE 4 ML WORKFLOW IN AGRICULTURE	22
FIGURE 5 COMPARATIVE TABLE OF DIFFERENT MACHINE LEARNING ALGORITHM IN CROP ANALYSIS	23
FIGURE 6 UAV IMAGE ANNOTATED WITH ML-BASED PREDICTIONS OF CROP HEALTH 24	
FIGURE 7 IMAGE SEGMENTATION IN CROPS USING U-NET.....	25
FIGURE 8 WEED SEGMENTATION USING U-NET	26
FIGURE 9 TRACKING TEMPORAL CHANGES WITH MEAN NDVI FOR GROWTH STAGES 28	
FIGURE 10 TRACKING NDVI IN TWO DIFFERENT REGIONS ACROSS TIME.....	30
FIGURE 11 VEGETATION INDICES WITH GROWING STAGES AND AGRONOMIC ACTIVITIES IN CROPS	31
FIGURE 12 WEATHER PATTERNS IN HIGHLANDS	33
FIGURE 13 WEATHER PATTERNS IN LOWLANDS	34
FIGURE 14 ELEVATION MAPS IN HIGHLAND AND LOWLAND	35
FIGURE 15 STUDY AREA IN GUATEMALA	37
FIGURE 16 IMAGE CAPTURE PREPARATION	38
FIGURE 17 EVALUATION CORN FIELDS HIGHLANDS	39
FIGURE 18 EVALUATION CORN FIELDS IN LOWLANDS	40
FIGURE 19 IMAGE CAPTURE SCHEDULE	41
FIGURE 20 SAMPLING SITES IN EACH LOCATION	42
FIGURE 21 TRAINING METRICS FOR EVALUATION, HIGHLANDS STAGE 1	51
FIGURE 22 TRAINING METRICS FOR EVALUATION, HIGHLANDS STAGE 2	52
FIGURE 23 TRAINING METRICS FOR EVALUATION, LOWLANDS STAGE 1	53
FIGURE 24 TRAINING METRICS FOR EVALUATION, LOWLANDS STAGE 2	53
FIGURE 25 CONFUSION MATRIX FOR EACH LOCATION AND EACH STAGE.....	54
FIGURE 26 VALIDATION METRICS, HIGHLANDS.....	55
FIGURE 27 VALIDATION METRICS, LOWLANDS STAGE 1 AND 2.....	57
FIGURE 28 NDVI DISTRIBUTION IN HIGHLANDS AND LOWLANDS	59
FIGURE 29 CROP GROWTH STAGES WITH STATISTICAL DISPERSION BY LOCATION.....	62

List of Tables

TABLE 1 COMPARATIVE TABLE OF REMOTE SENSING TECHNOLOGIES (SATELLITES, UAVS, HYPERSPECTRAL)	13
TABLE 2 U-NET VARIANTS	21
TABLE 3 COMPARISON OF KEY ENVIRONMENTAL FEATURES	34
TABLE 4 DESCRIPTION OF TILES GENERATION	43
TABLE 5 DESCRIPTION OF MASK GENERATION	44
TABLE 6 DATASET DISTRIBUTION FOR U-NET ANALYSIS	45
TABLE 7 HYPERPARAMETERS IN TRAINING LOOP, U-NET	45
TABLE 8 METRICS FOR VALIDATION ASSESSMENT, U-NET	47
TABLE 9 SHAPIRO - WILK TEST VARIABLES	48
TABLE 10 T-TEST VARIABLES	49
TABLE 11 PREDICTIONS VS GROUND TRUTH, HIGHLANDS	56
TABLE 12 PREDICTIONS VS GROUND TRUTH, LOWLANDS	58

Abbreviations

- **AI:** Artificial Intelligence
- **CIAT:** International Center for Tropical Agriculture (the acronym comes from Spanish)
- **CNN:** Convolutional Neural Network
- **CSV:** Comma – Separated Values
- **GIS:** Geographic Information System
- **IoU:** Intersection Over Union
- **JSON:** JavaScript Object Notation
- **FN:** False Negative
- **FP:** False Positive
- **ML:** Machine Learning
- **NDVI:** Normalized Difference Vegetation Index
- **NIR:** Near – Infrared
- **RGB:** Red, Green, Blue
- **TIFF:** Tagged Image File Format
- **TN:** True Negative
- **TP:** True Positive
- **U-Net:** Fully Convolutional Neural Network for Image Segmentation
- **UAV:** Unmanned Aerial Vehicle

1. Introduction

Sustainable agricultural practices today prioritize enhancing crop yields while optimizing agronomic activities. Unmanned Aerial Vehicles (UAVs), as a remote sensing technology, have revolutionized crop monitoring by providing high-resolution, multispectral imagery that captures essential vegetation indices such as the Normalized Difference Vegetation Index (NDVI) (Tucker, 1978). These indices are closely associated with crop vigor, biomass, and productivity, making them indispensable tools for precision agriculture (Zhang & Kovacs, 2012).

However, in regions like Guatemala, a significant number of smallholder farmers face challenges in accessing advanced agricultural technologies. Smallholder farmers make up over 80 percent of Guatemalan farmers, but in a country dominated by export-driven agriculture, they face limited access to land, financial tools, and markets (World Food Program, 2025).

Machine learning, especially Convolutional Neural Networks (CNNs), has become a cornerstone in agricultural analysis due to its ability to process complex image data efficiently (Zhang, et al., 2019). Among these, U-Net has emerged as a leading deep learning model for semantic segmentation, enabling precise delineation of crop fields from UAV imagery (Ronneberger, Fischer, & Brox, 2015).

Additionally, CNN-based techniques facilitate advanced applications like detecting nutrient deficiencies, pest infestations, and water stress, enabling timely interventions (Kamilaris & Prenafeta-Boldú, 2018). These types of methodologies enhance UAV imagery predictive power by integrating it with agronomic data and environmental factors, providing actionable insights for decision-making.

This study leverages UAV imagery and machine learning techniques to improve maize growth monitoring in Guatemala, focusing on two contrasting regions. By employing the U-Net deep learning architecture, we aim to delineate crop fields and extract NDVI metrics across critical growth stages (Ronneberger, Fischer, & Brox, 2015). This approach enables a deeper understanding of spatial variability in crop health, improving data-driven resource management.

This study is part of a project from **One CGIAR initiative and AGRILAC Resiliente¹**, where sustainable agricultural practices are promoted across Latin America and the Caribbean, addressing resilience to climate change and food security challenges (CIAT, 2023). This research contributes to the United Nations Sustainable Development Goals (SDGs), particularly SDG 2 (Zero Hunger), SDG 13 (Climate Action), and SDG 15 (Life on Land) (NATIONS, 2024), by advancing innovative solutions for sustainable farming.

Unlike traditional methods, this study focuses exclusively on UAV-derived data, emphasizing its superior spatial resolution and adaptability over satellite imagery (Persello & Bruzzone, 2010). The results are expected to provide actionable insights into maize growth dynamics, contributing to the broader discourse on agricultural resilience and sustainability in regions with diverse environmental conditions.

¹ This Initiative aims to increase the resilience, sustainability and competitiveness of Latin American and Caribbean agrifood systems and actors by better equipping them to meet urgent food security needs, reduce climate threats, stabilize conflict-vulnerable communities and reduce out-migration. (CGIAR, 2025)

2. Objectives

2.1 General

To characterize maize growth stages using UAV-derived NDVI imagery and machine learning techniques, and to relate these stages to agronomic yield data in contrasting environments (highlands and lowlands) in Guatemala.

2.2 Specifics

- To implement a U-Net-based segmentation framework for extracting crop fields from multispectral UAV imagery.
- To calculate NDVI-based metrics (mean, minimum, maximum, and standard deviation) for two maize growth phases and evaluate their spatial variability across regions.
- To analyze the relationship between NDVI statistics from each growth phase and agronomic yield data in Zacapa and Totonicapán.

3. Scientific overview

3.1 Remote Sensing in Agriculture

Overview of Remote Sensing Technologies and Their Role in Modern Agriculture

Remote sensing technologies have become indispensable in modern agriculture, offering a range of tools for monitoring crop health, optimizing resource allocation, and enhancing productivity. These technologies include satellite imagery, unmanned aerial vehicles (UAVs), and hyperspectral sensors, each providing unique benefits for agricultural management (Zhang & Kovacs, 2012). By capturing data across vast areas and at multiple temporal scales, remote sensing enables farmers and researchers to make informed decisions regarding irrigation, fertilization, and pest control, thereby increasing yields while reducing waste (Tao, et al., 2020).

Satellite systems, such as Sentinel-2 and Landsat, are widely used for large-scale agricultural monitoring. These systems provide multispectral data that can reveal vegetation health and land cover changes. However, their resolution and revisit times may limit their applicability for smallholder farms or time-sensitive interventions (Persello & Bruzzone, 2010). On the other hand, UAVs offer high-resolution imagery and can be deployed flexibly in time and area, making them ideal for precision agriculture (Li et al., 2021). In addition, hyperspectral sensors further enhance remote sensing capabilities by capturing detailed spectral signatures, allowing for advanced crop classification and stress detection (Zhang, et al., 2019).

Technology	Strengths	Limitations	Applications
Satellite Imagery	<ul style="list-style-type: none"> -Large-scale coverage -Frequent revisits (e.g., Sentinel-2 every 5 days) -Cost-effective for regional analysis 	<ul style="list-style-type: none"> -Lower spatial resolution (e.g., 10–30m for multispectral data) -Susceptible to cloud cover -Fixed acquisition times 	<ul style="list-style-type: none"> --Land cover classification -Climate and vegetation trend monitoring -Drought and soil moisture monitoring
UAV Imagery	<ul style="list-style-type: none"> -High spatial resolution (sub-meter level) -On-demand data collection -Operates below cloud cover -Customizable sensors 	<ul style="list-style-type: none"> -Limited area coverage per flight -Requires field deployment -Higher operational costs for large-scale monitoring 	<ul style="list-style-type: none"> -Precision agriculture and crop health assessment -Site-specific nutrient and water management -Crop damage detection

Hyperspectral Imagery	<ul style="list-style-type: none"> -Extremely detailed spectral information -Detects subtle differences in plant stress and disease -Suitable for advanced classification tasks 	<ul style="list-style-type: none"> - High data storage and processing requirements -Expensive sensors -Requires expert interpretation 	<ul style="list-style-type: none"> -Species-level vegetation classification -Early stress detection in crops -Disease and pest identification
-----------------------	--	--	--

Table 1 Comparative table of remote sensing technologies (satellites, UAVs, hyperspectral) (Zhang, et al., 2019)

Remote sensing also contributes to achieving global sustainable development goals (SDGs). For instance, SDG 2 (Zero Hunger) benefits from increased agricultural efficiency, while SDG 13 (Climate Action) is supported through the monitoring of crop responses to climatic changes. Initiatives like AgriLAC Resiliente leverage these technologies to promote resilience and innovation in agriculture across Latin America and the Caribbean.

Advantages of UAV Imagery Over Satellite-Based Systems

UAV imagery offers distinct advantages over satellite-based systems, particularly in precision, operational flexibility, and cost-effectiveness. These attributes have made UAVs a preferred choice for smallholder farms and regions with complex topographies (Zhang & Kovacs, 2012).

- High Spatial Resolution:** UAVs provide precise and adjustable spatial resolution, often achieving sub-meter levels, which is critical for identifying fine-scale agricultural features. This is particularly beneficial for farms where field sizes are often smaller than the resolution of many satellite systems (Zhang & Kovacs, 2012). For instance, Sentinel-2 has a spatial resolution of 10 meters for its multispectral bands, which may overlook critical intra-field variations (Persello & Bruzzone, 2010).
- Temporal Flexibility:** UAVs can be deployed on demand, ensuring data collection during critical crop growth stages. Unlike satellites, which are limited by fixed orbital schedules and potential cloud cover interference, UAVs offer the flexibility to operate under optimal weather conditions and at user-defined intervals (Zuo & Li, 2024). This capability is especially useful in dynamic environments where frequent monitoring is necessary.
- Enhanced Data Integration:** UAV data can be seamlessly integrated with other data sources, such as ground truth measurements and machine learning models, to provide

a comprehensive view of agricultural conditions. This integration enables better decision-making for resource optimization and yield prediction (Tao, et al., 2020).

- d. **Cost-Effectiveness for Localized Applications:** While satellite imagery is cost-efficient for large-scale monitoring, UAVs are more economical for small to medium-sized areas (Zuo & Li, 2024).
- e. **Customizable Sensor Payloads:** UAVs can carry a variety of sensors, including RGB cameras, multispectral cameras, and hyperspectral sensors, allowing users to customize data collection based on specific agricultural needs. For instance, hyperspectral sensors on UAVs can capture detailed spectral data for precise crop health assessment, which is often not feasible with standard satellite platforms (Zhang, et al., 2019).
- f. **Reduced Data Gaps and Cloud Interference:** Unlike satellite systems that may experience significant data gaps due to cloud cover or long revisit times, UAVs can operate below cloud levels and provide consistent data. This is particularly advantageous in tropical regions, where cloud cover frequently disrupts satellite observations.

Examples of NDVI Usage in Crop Monitoring and Growth Characterization

Integrating remote sensing technologies into corn cultivation can significantly enhance yield predictions and management practices. For instance, a study by Shanahan et al. (2001) demonstrated that the Green Normalized Difference Vegetation Index (GNDVI), derived from remote sensing imagery during mid-grain filling, exhibited strong correlations with grain yield, with coefficients reaching up to 0.95 in certain years. This high correlation suggests that remote sensing can serve as a reliable proxy for estimating crop yield variability within fields. While specific percentage increases in yield due to the adoption of remote sensing tools can vary based on factors such as environmental conditions, management practices, and the specific technologies employed, the ability to monitor crop health and development dynamically allows for more informed decision-making.

The Normalized Difference Vegetation Index (NDVI) has proven to be a vital metric in monitoring and characterizing crop growth, first because of its ability to provide insights into vegetation health through its sensitivity to chlorophyll content and biomass levels. Derived from red and near-infrared (NIR) spectral bands, NDVI measures the differential reflectance

of vegetation, making it a reliable indicator for assessing plant vigor. In agricultural applications, NDVI plays a critical role in tracking phenological changes, as it captures the progression of growth stages from emergence to maturity. For instance, studies have highlighted its effectiveness in monitoring maize and wheat crops, where temporal NDVI patterns closely correlate with biomass accumulation and leaf area index (Tao, et al., 2020).

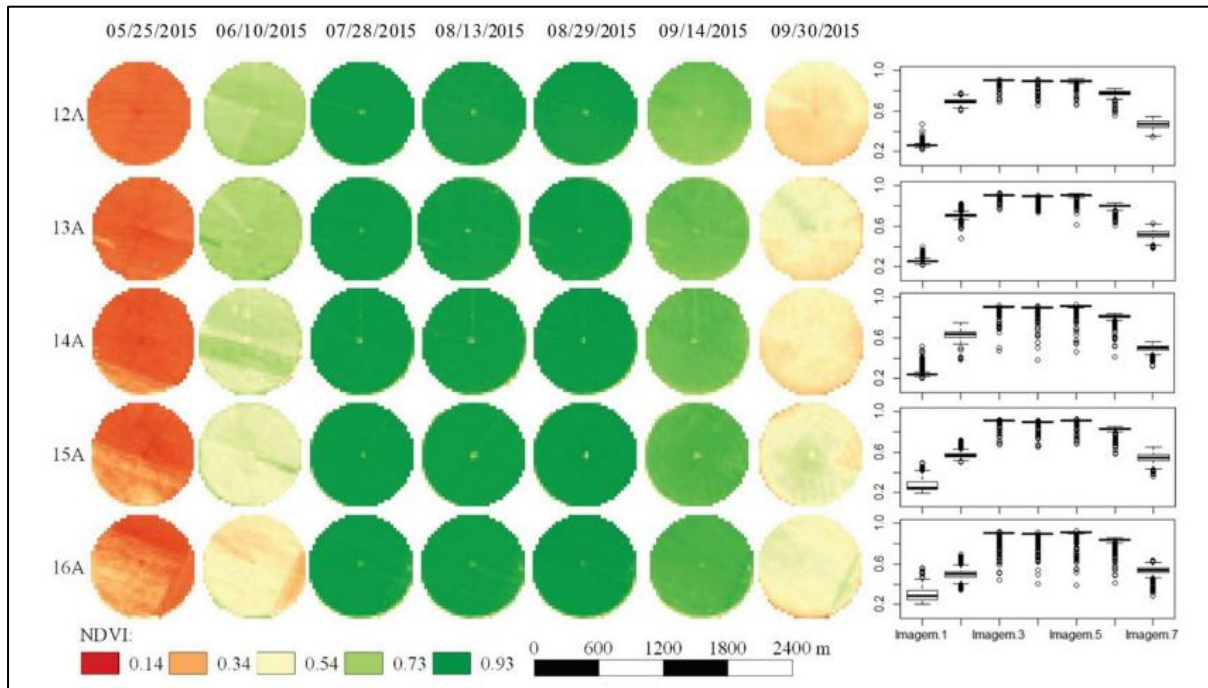


Figure 1 NDVI across crop growth stages (Alvino, Alemán, Filgueiras, Althoff, & Cunha, 2020)

Stress detection is another crucial application of NDVI. Crops subjected to water deficits, nutrient imbalances, or pest infestations exhibit changes in chlorophyll levels that NDVI can detect early, allowing for timely intervention. This capability reduces the risk of yield losses and optimizes resource use, as targeted fertilization can be applied based on identified areas of stress. Furthermore, NDVI facilitates precision agriculture practices by mapping spatial variability within fields. This spatial data helps farmers adjust input levels, such as fertilizers and pesticides, to specific field zones, thereby minimizing waste and environmental impact (Zuo & Li, 2024)

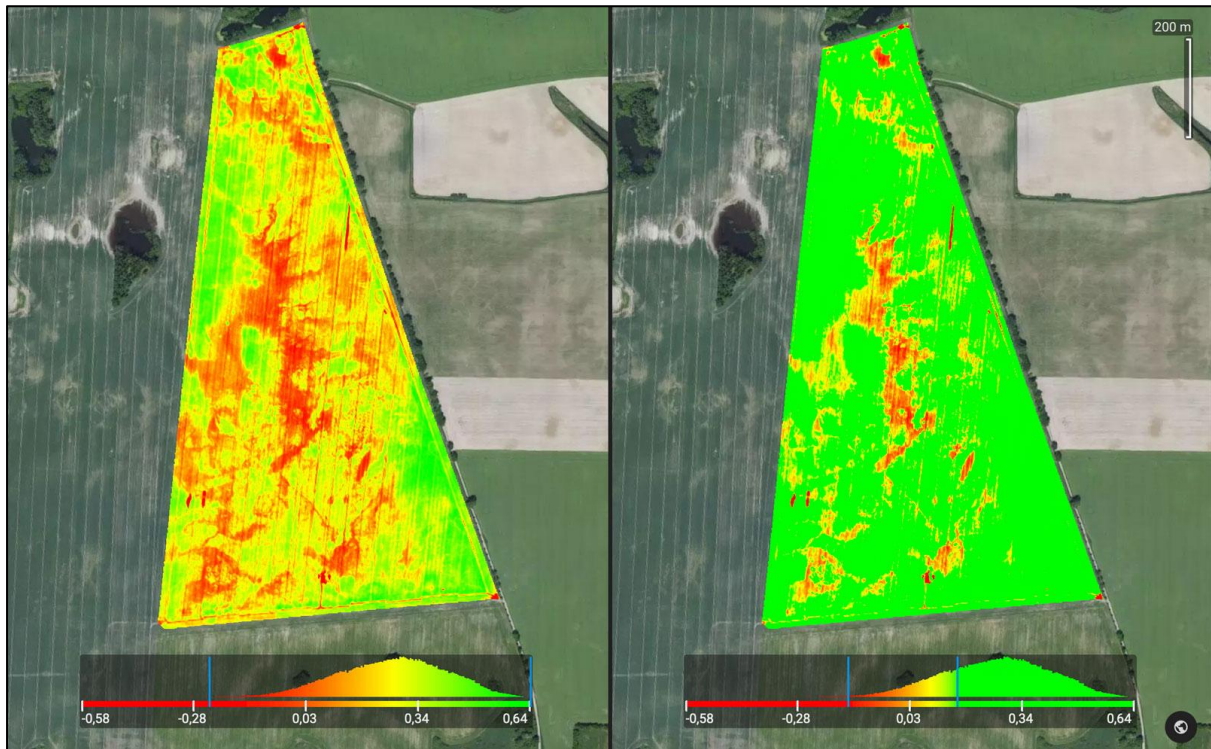


Figure 2 Crop variability using NDVI from high-resolution UAV image (PIX4D, 2025)

According with Zhang & He, 2019, recent research in remote sensing and machine learning have expanded the utility of NDVI in agriculture. By integrating NDVI data with predictive models, researchers can estimate crop yields with greater accuracy, even under varying environmental conditions. For example, NDVI trends over time have been used to forecast yields by correlating vegetation vigor with historical productivity data. Additionally, the combination of NDVI with other indices, such as the Enhanced Vegetation Index (EVI), further refines assessments of vegetation health and stress.

3.2 Machine Learning in Crop Analysis

Machine Learning and Neural Networks for Image Processing

Machine Learning (ML) has made significant advances on image processing fields by enabling automated feature extraction, pattern recognition, and decision-making. Unlike traditional algorithms that rely on manually defined features, ML models, particularly **deep learning**, learn to identify relevant patterns directly from data. This capability has made deep

learning the standard approach for complex image-related tasks such as object detection, segmentation, and classification (LeCun, Bengio, & Hinton, 2015).

One of the most widely adopted deep learning architectures for image analysis is the **Convolutional Neural Network (CNN)**. CNNs apply multiple layers of convolutional filters to extract spatial hierarchies of features—starting with simple edges and progressing to complex patterns such as textures, shapes, and objects. These networks reduce the computational burden by utilizing shared weights and local receptive fields, making them particularly effective for image processing compared to fully connected neural networks (Cheng, Papandreou, Kokkinos, Murphy, & Yuille, 2018).

Types of Neural Networks and Their Applications

Various neural network architectures address different ML challenges:

- **Feedforward Neural Networks (FNNs):** The simplest form of artificial neural networks, where data moves in one direction without loops, primarily used for classification tasks.
- **Convolutional Neural Networks (CNNs):** Specifically designed for image analysis, CNNs automatically extract spatial features at multiple hierarchical levels from visual data.
- **Recurrent Neural Networks (RNNs):** Used for sequential data processing, particularly in speech recognition, time-series forecasting, and natural language processing.
- **Autoencoders:** Applied in unsupervised learning for dimensionality reduction, denoising, and anomaly detection.
- **Generative Adversarial Networks (GANs):** A framework where two networks (generator and discriminator) compete to produce realistic synthetic data, widely used in image generation and enhancement.

Deep learning frameworks have evolved to include transformers for vision-based tasks, such as Vision Transformers (ViTs) and SegFormers, which leverage self-attention mechanisms to

model long-range dependencies. While these newer models are gaining popularity, CNNs remain dominant in image segmentation due to their computational efficiency and strong feature extraction capabilities (LeCun, Bengio, & Hinton, 2015).

Comparing Image Segmentation Models

Image segmentation is a crucial ML task that involves classifying every pixel in an image. Various deep learning architectures have been developed to address segmentation problems:

- **Fully Convolutional Networks (FCNs):** A modified CNN architecture that removes fully connected layers, enabling dense per-pixel classification.
- **DeepLabV3+:** Introduces “atrous convolution” to extract multi-scale features, improving segmentation accuracy for objects with varying sizes.
- **Mask R-CNN:** A two-stage model that performs both object detection and segmentation, excelling in instance segmentation tasks.
- **Transformers for Vision (SegFormer, Swin Transformer):** Modern architectures that use self-attention to capture contextual information across an entire image, improving segmentation quality.

Although these architectures provide state-of-the-art results, U-Net remains one of the most widely used segmentation models due to its balance between computational efficiency and accuracy, especially in high-resolution image analysis.

U-Net: A Robust Model for Image Segmentation

U-Net is a fully convolutional neural network (FCN) designed for pixel-wise classification, making it ideal for segmentation tasks. Initially introduced for biomedical image analysis (Ronneberger, Fischer, & Brox, 2015), its encoder-decoder architecture with skip connections allows it to outperform conventional segmentation models in various domains. The network’s ability to learn from small datasets while preserving spatial details has contributed to its widespread adoption.

Architecture of U-Net

U-Net follows a **symmetric encoder-decoder structure**:

- **Encoder (Contracting Path):** A series of convolutional layers and max-pooling operations that progressively reduce image dimensions while increasing feature depth. This process helps capture high-level abstract features.
- **Bottleneck:** The transition layer between the encoder and decoder, where features are processed at their lowest spatial resolution.
- **Decoder (Expanding Path):** Utilizes up-convolutions (transposed convolutions) to restore spatial resolution and reconstruct segmented regions.
- **Skip Connections:** Directly link corresponding encoder and decoder layers to preserve fine-grained spatial details, which would otherwise be lost due to pooling.

This design makes U-Net highly effective at segmenting objects with irregular shapes and complex boundaries. Unlike traditional segmentation models that rely heavily on **fully connected layers**, U-Net leverages only convolutional operations, making it computationally efficient.

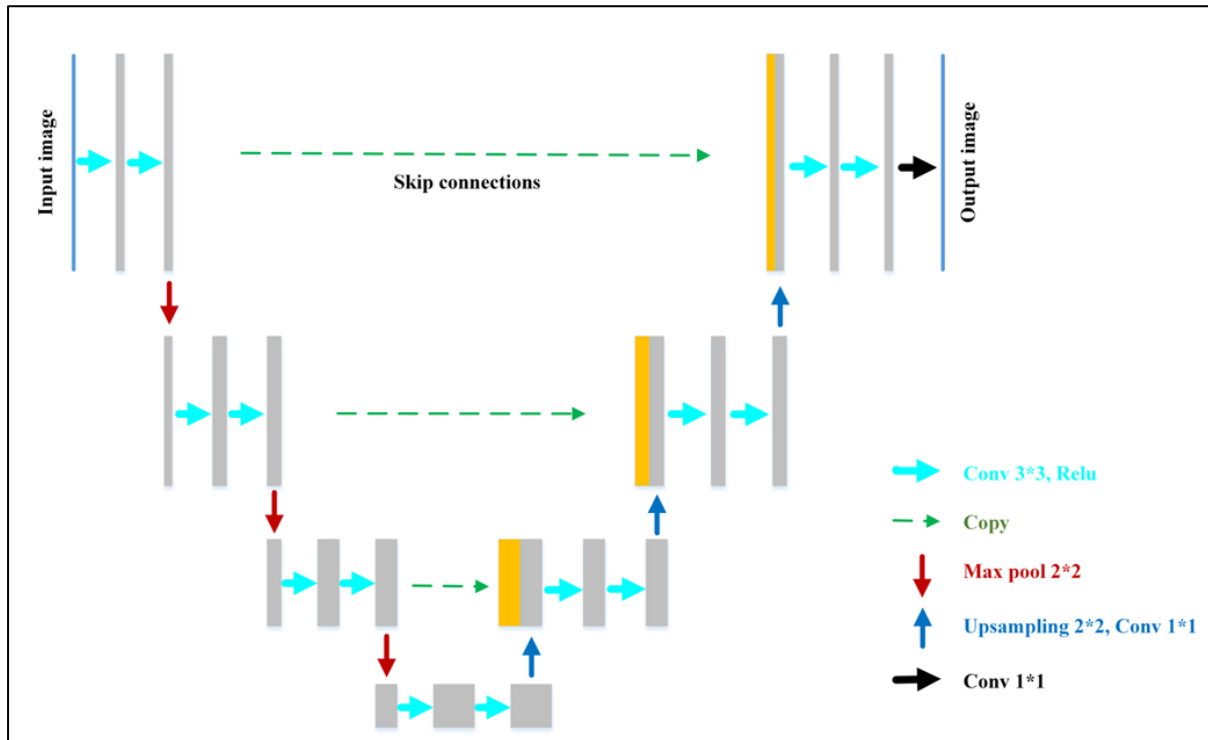


Figure 3 U-Net architecture diagram (Sui, Jiang, Zhang, & Pan, 2020)

Variants and Enhancements of U-Net

To improve its segmentation capabilities, various U-Net modifications have been proposed:

- **Attention U-Net:** Introduces attention gates to highlight relevant image regions while suppressing background noise (Oktay, et al., 2018).
- **ResUNet:** Incorporates residual connections, allowing deeper networks to learn efficiently and avoid gradient vanishing (Zhang & Liu, 2018).
- **Nested U-Net (UNet++):** Uses **dense connections** to refine feature maps at multiple levels, improving segmentation accuracy .
- **Lightweight U-Net:** Designed for real-time applications by reducing the number of parameters while maintaining performance (Wang, Zhang, Wang, Ma, & Lui, 2020).

Each of these U-Net variants improves segmentation quality depending on the dataset characteristics, computational constraints, and real-world application needs.

Variant	Key Improvements	Applications	Reference
Attention U-Net	Integrates attention mechanisms to focus on relevant features, enhancing segmentation accuracy in complex images.	Medical image segmentation where distinguishing fine details is crucial.	(Oktay, et al., 2018)
ResU-Net	Incorporates residual connections to facilitate deeper network training and mitigate vanishing gradient issues.	Biomedical image analysis requiring deep feature extraction.	(Zhang & Liu, 2018)
UNet++	Features nested and dense skip connections to improve gradient flow and capture multi-scale features.	Segmentation tasks demanding high accuracy and detail.	(Zhou, Siddiquee, Tajbakhsh, & Liang, 2018)
3D U-Net	Extends U-Net to three dimensions for volumetric data, enabling context capture in 3D space.	3D medical imaging modalities like MRI and CT scans.	(Cicek, Abdulkadir, Lienkamp, Brox, & Ronneberger, 2016)
R2U-Net	Combines residual and recurrent connections to model contextual information and improve segmentation performance.	Applications requiring context-aware segmentation.	(Alom, Hasan, Yakopcic, Taha, & Asari, 2018)

Table 2 U-Net variants

Strengths and Limitations of U-Net

Advantages:

- High segmentation accuracy due to skip connections preserving spatial resolution.
- Efficient learning with limited data, making it useful in scenarios with small labelled datasets.
- Flexible architecture that can be extended to various domains with minimal modifications.
- Computational efficiency compared to transformer-based models.

Limitations:

- **Loss of fine-grained features in deeper layers**, which is partially mitigated by nested U-Net structures.

- **Higher memory requirements due to extensive feature maps**, making training slower on large images.
- **Limited robustness to noisy datasets**, which can lead to misclassifications in cluttered backgrounds.

Introduction to Machine Learning Techniques in Agriculture

Machine learning (ML) is leading agricultural innovation, transforming data into actionable insights through advanced computational mechanism. ML enables predictive analysis, automates complex processes, and enhances the precision of resource allocation in agricultural management. Among the various branches of ML, deep learning has emerged as a transformative tool due to its capacity to analyze high-dimensional data and uncover non-linear relationships that traditional statistical methods often miss (Zhang, et al., 2019).

Deep learning techniques, particularly convolutional neural networks (CNNs), have proven instrumental in agricultural applications such as crop classification, pest detection, and disease diagnosis. These models excel in image-based analysis by leveraging their ability to extract hierarchical features from agricultural imagery. For instance, CNNs have been utilized to differentiate healthy crops from diseased ones, even under challenging conditions involving occlusion or poor lighting.

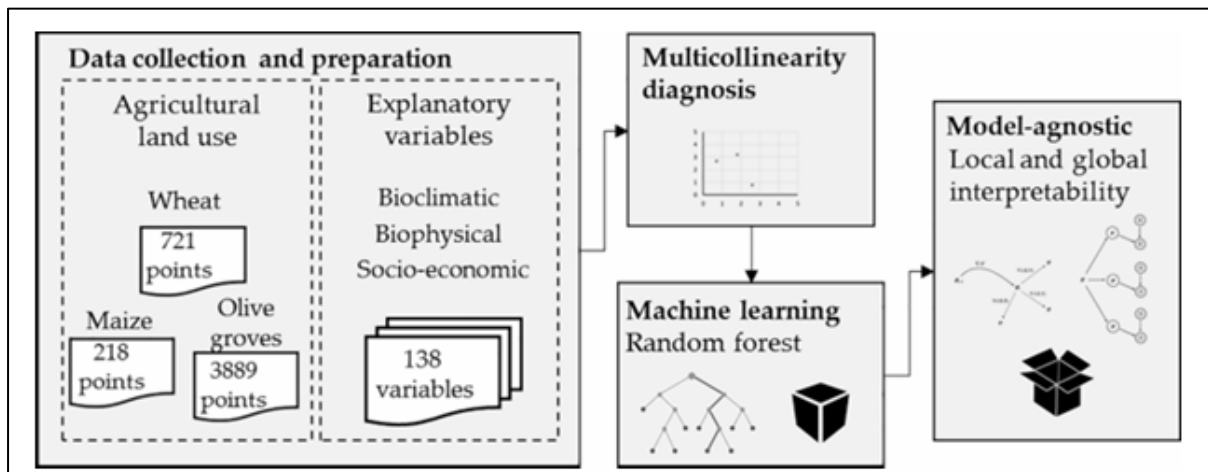


Figure 4 ML workflow in agriculture (Viana, Freire, Santos, & Abrantes, 2021)

Another critical aspect of ML in agriculture is its integration with multi-source data, including UAV imagery, satellite datasets, and ground-based sensors. By combining these inputs, ML

models can create comprehensive analyses that support precision farming practices. For example, models trained on UAV data have demonstrated high accuracy in detecting intra-field variability, enabling targeted interventions that optimize resource use while minimizing environmental impact (Tao, et al., 2020).

Algorithm	Accuracy	Strengths	Limitations
CNN (Convolutional Neural Networks)	95–99%	High accuracy in image classification, effective for crop disease detection.	Requires large datasets and high computational resources.
RF (Random Forest)	80–98%	Robust for crop classification, handles mixed data types well.	Sensitive to noisy data and high dimensionality.
Support Vector Machines (SVM)	85–97%	Good for high-resolution image classification, works well with small datasets.	Computationally expensive with large datasets, sensitive to hyperparameters.

Figure 5 Comparative table of different Machine Learning algorithm in crop analysis (Persello & Bruzzone, 2010)

Emerging methodologies, such as transfer learning and unsupervised learning, further enhance the applicability of ML in agriculture. Transfer learning allows pre-trained models to adapt to new datasets with limited samples, significantly reducing the time and computational resources required for training. Unsupervised learning, on the other hand, identifies hidden patterns in data without relying on labelled examples, offering insights into complex agricultural systems.

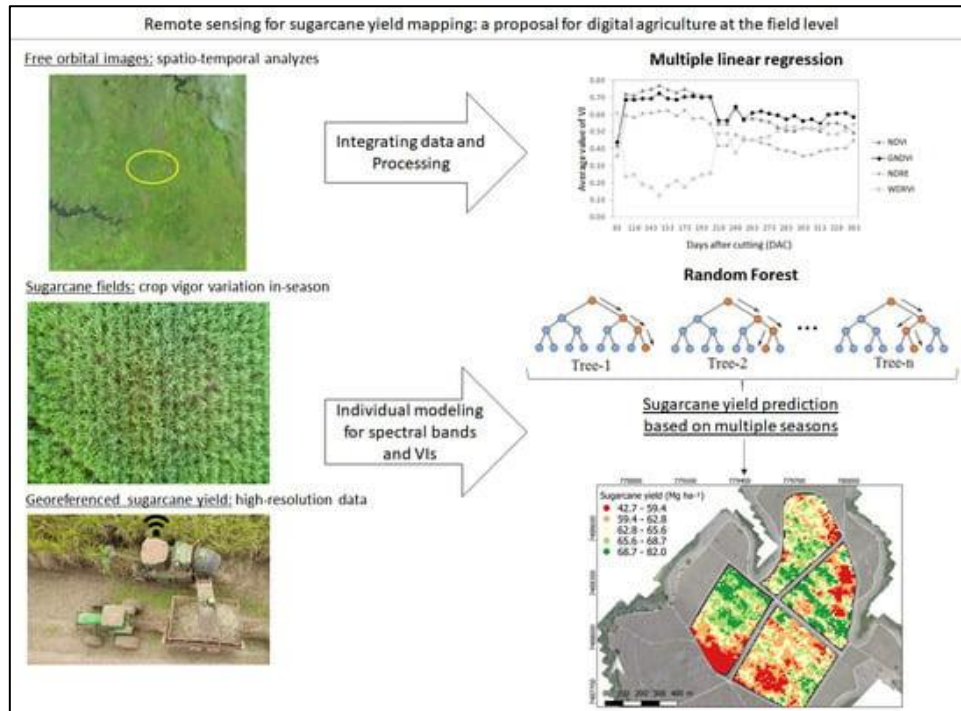


Figure 6 UAV image annotated with ML-based predictions of crop health (Canata, Wei, Maldaner, & Molin, 2021)

Applications of U-Net in Delineating Areas of Interest, Specifically Crop Fields

The U-Net architecture has led numerous applications in the agricultural sector, transforming how crop fields are monitored, analyzed, and managed. U-Net has been widely adopted for segmenting crop fields from background elements such as soil, weeds, and non-cultivated areas. By generating high-quality segmentation masks, it enables researchers and farmers to isolate crop-specific regions for detailed analysis. For instance, UAV imagery processed with U-Net can identify and classify individual crop rows, detect gaps due to poor germination, and distinguish between different crop types. This level of granularity is essential for precision agriculture, where decisions must be made at the intra-field level (Ronneberger, Fischer, & Brox, 2015)

In addition to its role in crop investigation, U-Net has been employed to detect and map areas affected by stressors such as pests, diseases, and water deficiency. By highlighting stressed regions, allowing for targeted interventions. For example, regions with lower chlorophyll levels, as indicated by vegetation indices like NDVI, can be accurately delineated, helping

farmers focus their resources on problem areas. This application not only optimizes input usage but also mitigates potential yield losses by enabling timely responses (Tao, et al., 2020).

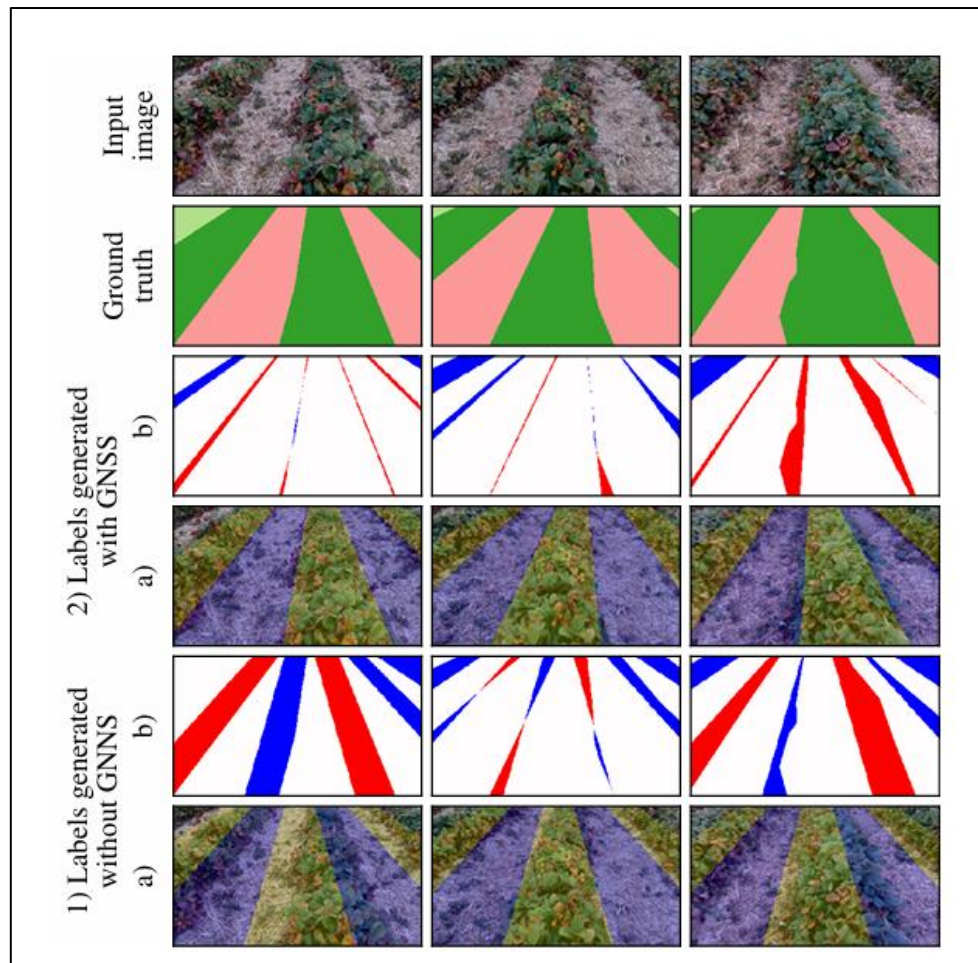


Figure 7 Image segmentation in crops using U-Net (Bakken, Ponnambalam, From, & Moore, 2021)

Another critical application of U-Net is in the segmentation of multispectral and hyperspectral data for large-scale agricultural monitoring. Its ability to handle multi-band data allows it to differentiate between crops and other vegetation types, providing insights into land use patterns and crop diversity. Supporting governmental and organizational efforts in agricultural planning, subsidy allocation, and compliance monitoring. For instance, segmentation maps generated by U-Net can validate crop insurance claims by accurately identifying field boundaries and crop types (Zhang, et al., 2019).

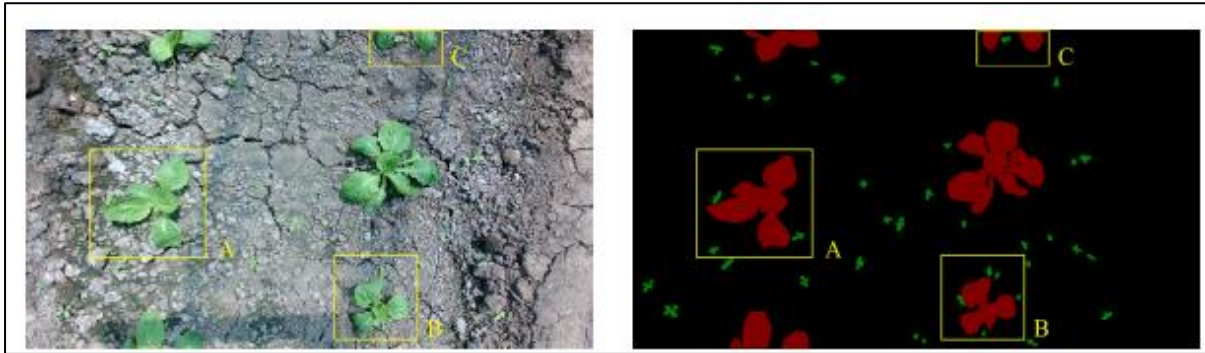


Figure 8 Weed segmentation using U-Net (Ma, et al., 2023)

Variants of U-Net, incorporating attention mechanisms and lightweight designs, have been successfully applied to rapidly changing environments, such as during different growth phases of crops. These adaptations ensure that the architecture remains effective even in scenarios with high temporal variability, such as monitoring phenological changes or seasonal planting patterns (Ronneberger, Fischer, & Brox, 2015).

3.3 Characterization of Growth Stages Using NDVI

Significance of Characterizing Growth Stages for Sustainable Agriculture

Characterizing crop growth stages is essential for optimizing agricultural practices and sustainable food production. By understanding the specific needs of crops during different growth phases, farmers can implement targeted interventions that improve resource efficiency, such as applying fertilizers or irrigation at the right time and in the adequate amounts. This precision not only help by improving yields but also minimizes the environmental impact of agricultural activities, aligning with sustainable development goals such as SDG 2 (Zero Hunger) and SDG 13 (Climate Action) (Zhang & Kovacs, 2012).

The identification of growth stages also enhances the ability to forecast yields and detect potential stressors early. During critical phases, such as flowering or grain filling, timely responses to stress can prevent significant yield losses. Moreover, characterizing growth stages supports better planning for harvesting and post-harvest activities, contributing to reduced food waste and improved supply chain efficiency (Tao, et al., 2020). On a larger scale, the ability to monitor crop growth stages across regions provides valuable data for agricultural policy-making and food security planning. By leveraging tools like NDVI and remote sensing,

stakeholders can gain insights into regional crop performance, enabling informed decisions on resource allocation and market strategies (Li, Wang, Wang, & Qi, 2024).

Explanation of NDVI Metrics in Assessing Crop Vigor and Growth

Normalized Difference Vegetation Index (NDVI) is a widely used metric for assessing crop vigor and growth due to its sensitivity to vegetation health and chlorophyll content. Derived from the red and near-infrared (NIR) bands (Blue and Green are often used instead of Red), NDVI values range from -1 to +1, with higher values indicating healthier vegetation. Several key NDVI metrics are utilized to characterize crop growth, each offering unique insights into plant performance (Tao, et al., 2020).

Example NDVI map with annotated zones for mean, maximum, and minimum values.

- **Mean NDVI:** Represents the average vegetation vigor across a field, providing an overview of overall crop health. Mean NDVI is particularly useful for comparing fields or regions and identifying large-scale trends in crop development (Li, Wang, Wang, & Qi, 2024)
- **Maximum NDVI:** Highlights areas of peak vegetation vigor, often corresponding to the healthiest plants. This metric can indicate optimal zones for seed production or harvesting (Zuo & Li, 2024).
- **Minimum NDVI:** Identifies areas with low vegetation vigor, which may signal stress factors such as pest infestation, water deficiency, or nutrient depletion. By locating these zones, farmers can target interventions more effectively (Zuo & Li, 2024)
- **Standard Deviation of NDVI:** Reflects spatial variability within a field, highlighting heterogeneity in crop performance. High variability may indicate uneven application of inputs or localized stressors, plants are growing in an unequal rate (Tao, et al., 2020)

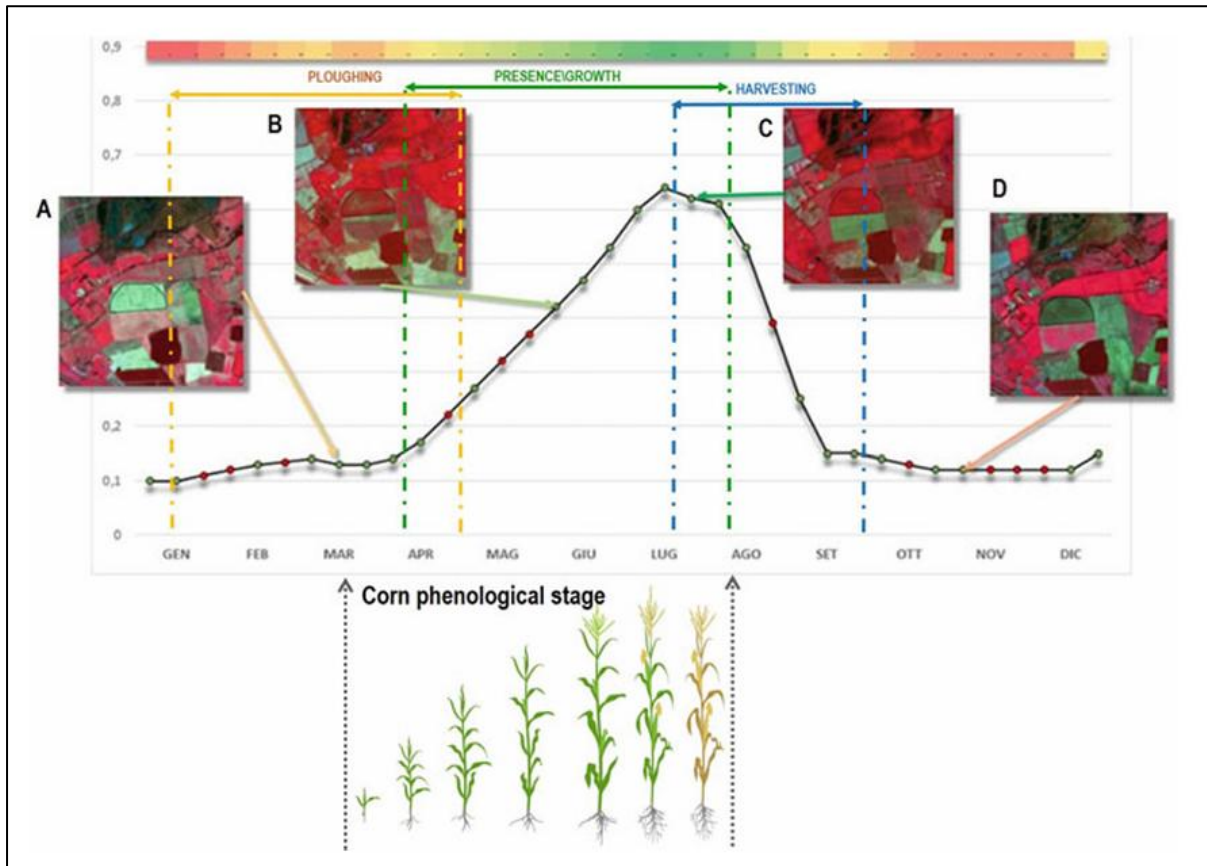


Figure 9 Tracking temporal changes with mean NDVI for growth stages (Copacenaru, et al., 2021)

In the image above, some examples about tracking crop stages (A: Plowing; B: Vegetative stage; C: Harvesting period; D: Area after harvesting) using images for further analysis using metrics mentioned. These metrics, when analyzed together, provide a comprehensive understanding of crop performance. For instance, comparing maximum and minimum NDVI values can reveal the degree of variability within a field, while temporal changes in mean NDVI can track growth dynamics over time.

Use of T-Test for Analyzing NDVI Differences Between Regions

The Independent Samples T-Test is a statistical method used to compare the means of two groups and determine if the difference between them is statistically significant. In the context of analyzing NDVI from corn fields in two contrasting regions in Guatemala, the T-test is particularly important for several reasons:

- **Objective Evaluation of Differences:** NDVI is a widely used indicator of vegetation health, productivity, and biomass. Differences in NDVI between regions can indicate variations in agricultural conditions such as soil fertility, water availability, climate, or farming practices. The T-test allows researchers to evaluate these differences quantitatively, ensuring conclusions are data-driven and unbiased (Fisher, 1925).
- **Significance of Results:** While visual inspection of NDVI values might suggest differences between Totonicapán and Zacapa, these observations could result from random variability rather than real disparities. The T-test uses statistical rigor to determine whether the observed differences are significant, reducing the risk of drawing incorrect conclusions (Montgomery, 2013)
- **Importance in Environmental and Agricultural Research:** In agricultural research, understanding spatial differences in NDVI helps identify factors driving regional disparities. For example, Totonicapán, with its cooler highland climate, may differ in vegetation health compared to Zacapa's warmer and drier conditions. Such findings can guide sustainable farming practices and climate-resilient agricultural policies.

By applying the T-test, an analysis not only strengthens the statistical validity of the conclusions but also provides actionable insights for improving agricultural outcomes and addressing environmental challenges in specific regions. The following image shows how NDVI differs from two close regions (purple rectangle).

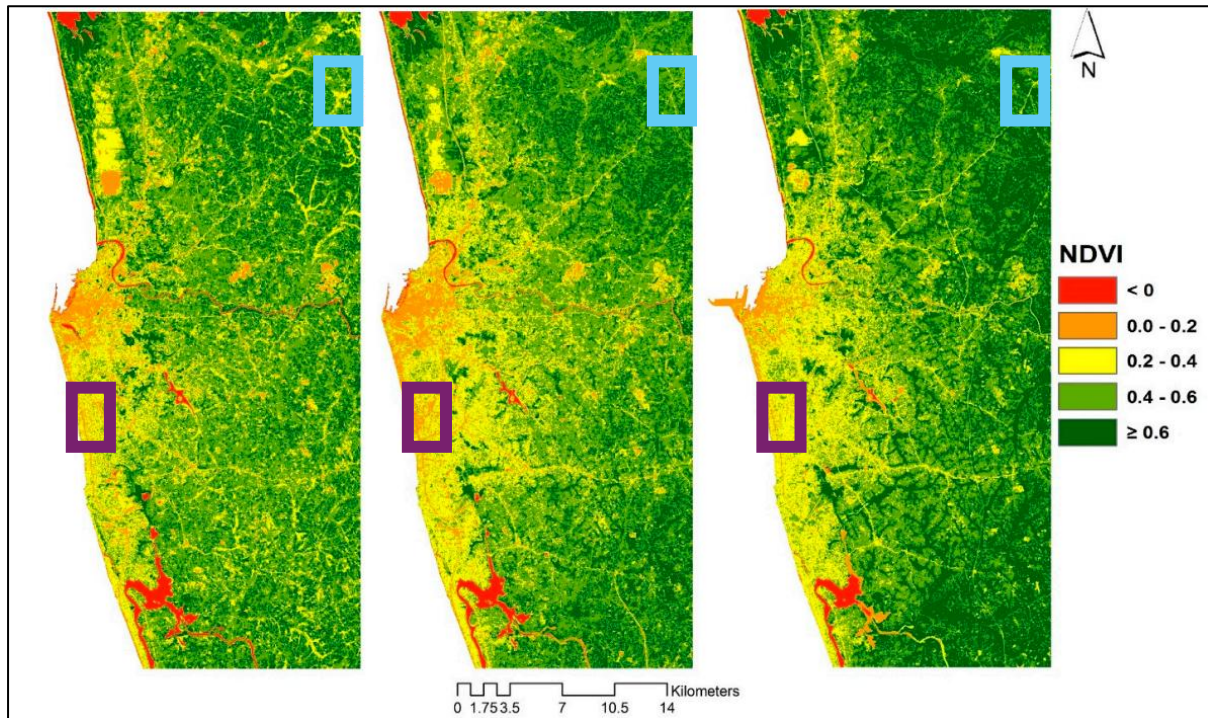


Figure 10 Tracking NDVI in two different regions across time (Ranagalage, Estoque, & Murayama, 2018)

3.4 Temporal and Spatial Analysis of NDVI in Phenological Changes

Temporal and spatial analysis of NDVI offers profound benefits in detecting phenological changes. By monitoring NDVI values over time, critical growth transitions, such as emergence, vegetative growth, flowering, and senescence (Chlorophyll decays, causing leaves to change color) can be identified. These transitions are essential for aligning agricultural practices with crop requirements (Zhang & Kovacs, 2012).

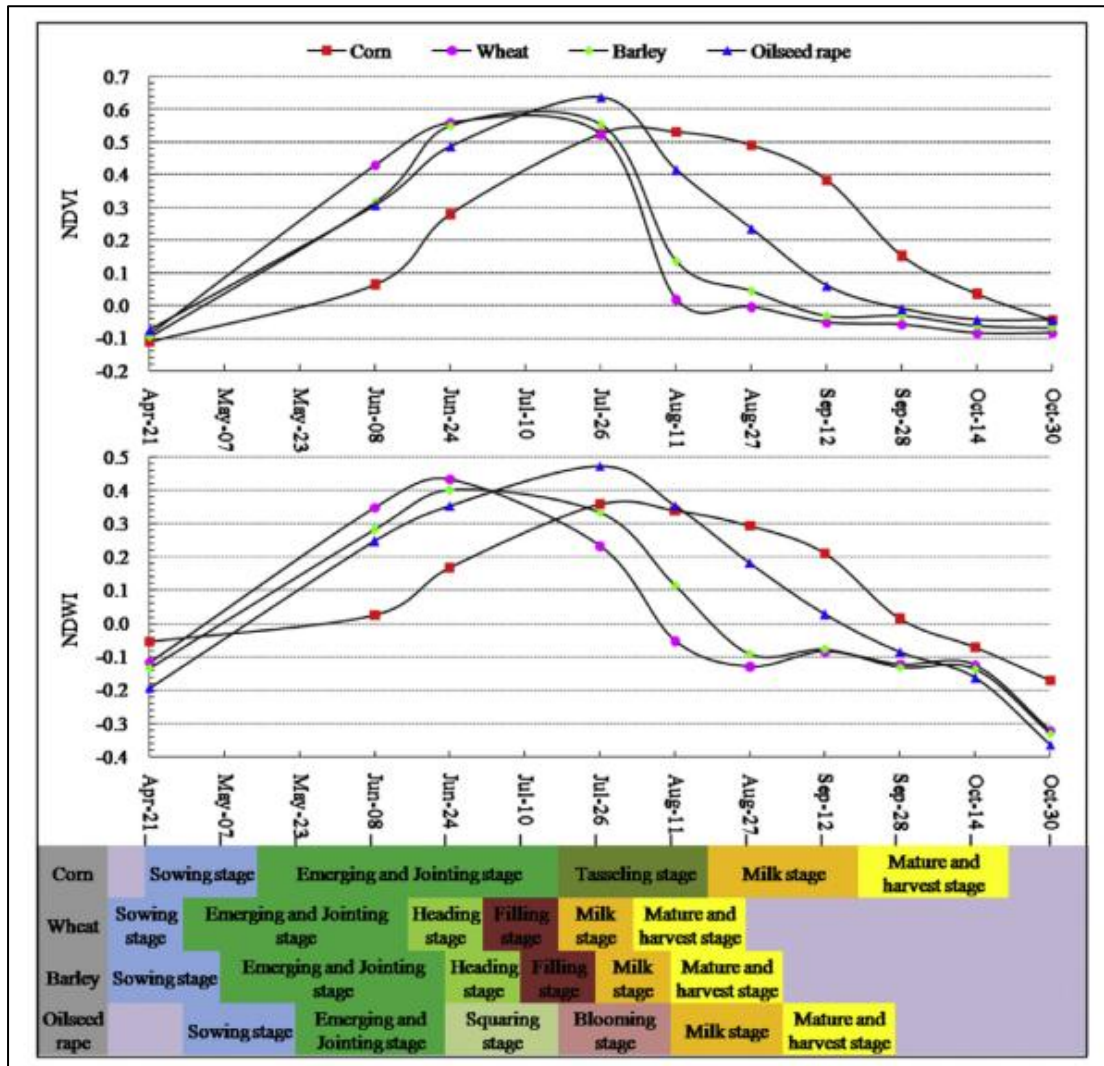


Figure 11 Vegetation indices with growing stages and agronomic activities in crops (Liu, Song, & Deng, 2016)

Spatial analysis complements temporal trends by revealing intra-field variability, which is often masked in field-wide averages. This analysis helps identify zones with differing growth patterns, supporting site-specific management. For instance, NDVI maps can delineate areas requiring additional attention, while avoiding over-application (Tao, et al., 2020).

Examples of Studies Linking Remote Sensing Indices with Yield Data for Decision-Making

Numerous studies have demonstrated the utility of integrating remote sensing indices like NDVI with yield data for improved agricultural decision-making. For instance, Tao et al. (2020) highlighted how UAV-based hyperspectral imagery and NDVI values were used to

predict maize yields in diverse environmental conditions. Their findings showed that integrating NDVI metrics, such as mean and standard deviation, with yield data improved the accuracy of crop performance predictions.

In another study, Zhang et al. (2019) used machine learning models to analyze the relationship between NDVI values and rice yields. By incorporating both remote sensing and agronomic data, their approach enabled the identification of key factors influencing productivity, such as irrigation (if applicable) patterns and nutrient levels. These insights supported the development of targeted management strategies, ultimately enhancing yield outcomes.

Additionally, Jones et al. (2021) conducted a study examining wheat yield predictions across different soil types in the United States using NDVI metrics. They demonstrated that early-season NDVI values, combined with soil moisture data, could reliably predict end-of-season yields. This approach highlighted the importance of integrating remote sensing with environmental data to address region-specific challenges in crop production.

These examples underline the importance of combining remote sensing and ground-based data to create robust frameworks for precision agriculture. By leveraging such integrated approaches, stakeholders can improve decision-making, optimize resource use, and achieve sustainable agricultural productivity.

3.5 Regional Context: Zacapa and Totonicapán

Environmental and Agronomic Characteristics

a. Totonicapán: Highlands with Favorable Conditions for Agriculture

Totonicapán is located in the western highlands of Guatemala, characterized by its temperate climate, fertile soils, and mountainous topography. These factors make it an ideal region for cultivating staple crops like maize and beans, which form the backbone of local agriculture.

- **Climate:** The region experiences mild temperatures, with an annual average ranging from 10° C to 22° C. Rainfall is evenly distributed throughout the year, averaging 1,200 mm annually. These conditions support the robust growth of crops, particularly during the vegetative and reproductive phases of maize.

- **Soils:** Andisols dominate Totonicapán, known for their high organic matter content, good moisture retention, and fertility. These attributes enable small-scale farmers to achieve consistent productivity with minimal external inputs.
- **Land Use:** Coniferous forests cover 41% of the area, while 30% is dedicated to crop production. Small-scale maize plots predominate, managed under traditional practices like using local seed varieties and maintaining soil fertility through crop rotation and organic inputs.

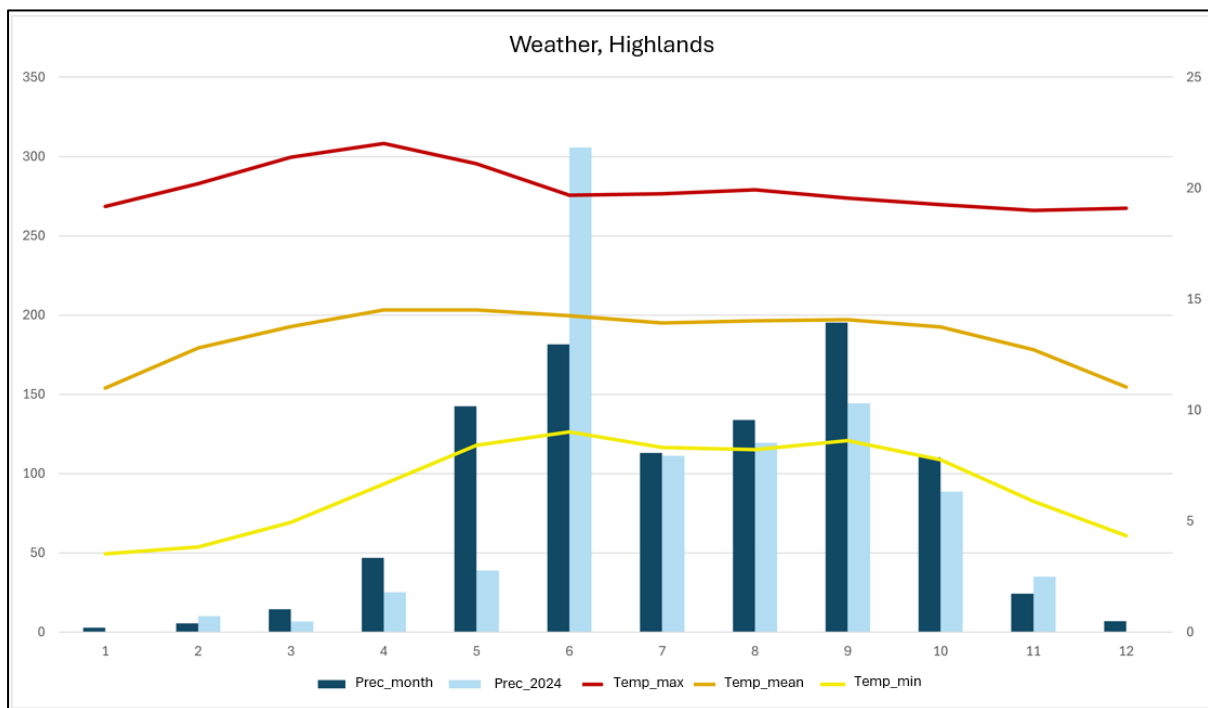


Figure 12 Weather patterns in Highlands (INSIVUMEH, 2023)

b. Zacapa: Lowlands with Suboptimal Agricultural Conditions

In contrast, Zacapa is situated in the Motagua Valley in eastern Guatemala. It presents a challenging agricultural environment due to its warmer and drier climate, less fertile soils, and lower elevation.

- **Climate:** Zacapa has a semi-arid climate with temperatures ranging from 18° C to 32°C. Annual rainfall averages around 800 mm, making water management a critical factor for sustaining agriculture. Some monitored areas, however, benefit from microclimates with higher precipitation due to their location in mountainous regions.

- **Soils:** Predominantly inceptisols with variable organic matter and low water retention capacity. These soils require intensive management to maintain productivity, including the use of fertilizers and supplemental irrigation.
- **Land Use:** The region's agricultural landscape includes a mix of short-cycle crops like maize and sorghum (7%), perennial crops such as coffee (25%), and shrub vegetation (40%). Maize cultivation is often restricted to smaller plots due to limited water availability.

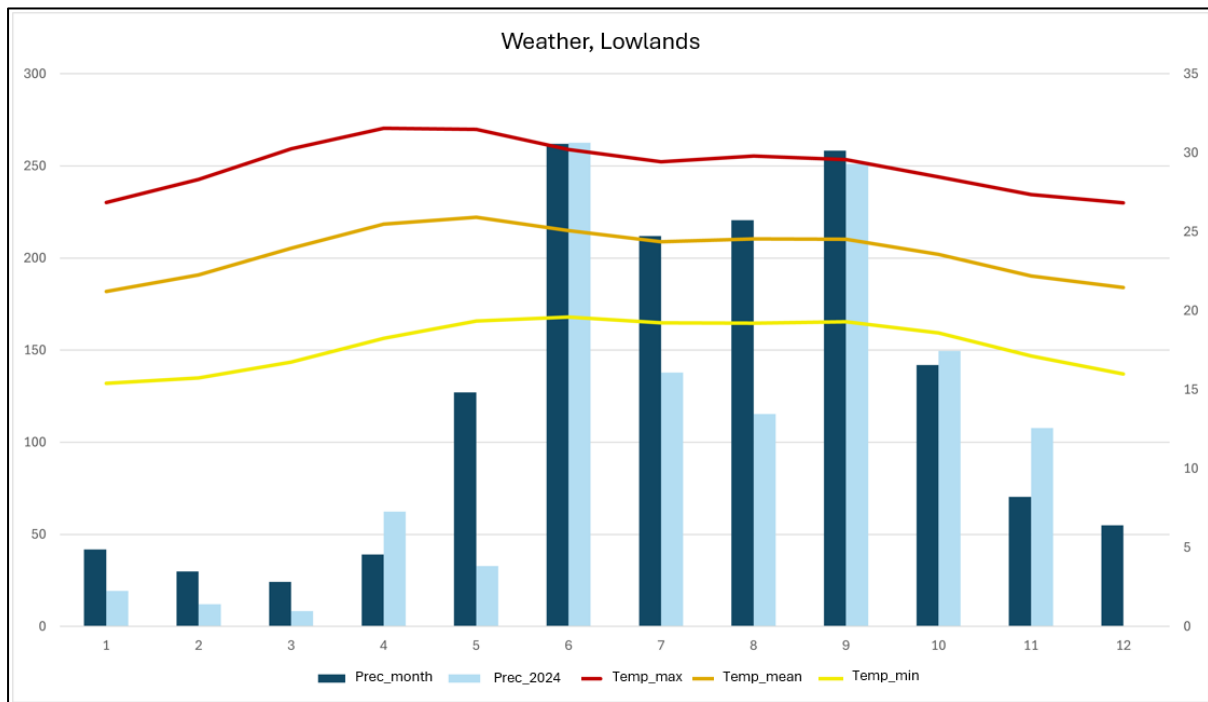


Figure 13 Weather patterns in Lowlands (INSIVUMEH, 2023)

Feature	Totonicapán	Zacapa
Elevation (m)	2,000-3,000	300-1,000
Annual Rainfall (mm)	~1,200	~800
Dominant Soil Type	Andisols: highly fertile, and ideal for agriculture.	Inceptisols: moderately fertile.
Temperature Range (°C)	10-22	18-32

Table 3 Comparison of Key Environmental Features (MAGA., 2019)

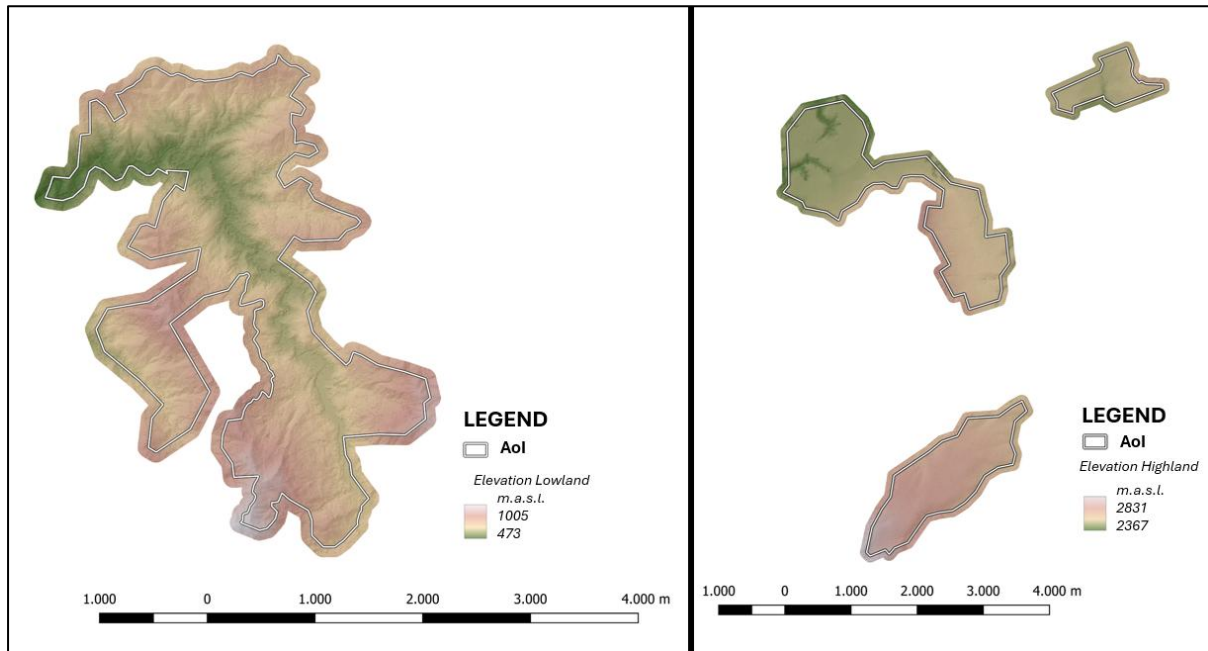


Figure 14 Elevation maps in highland and lowland (Martínez, Chávez-Can, & Navarro, 2024)

c. Agronomic Practices

Totonicapán and Zacapa differ significantly in terms of agricultural management due to their contrasting environmental conditions.

Totonicapán:

- Farmers rely on traditional techniques, including the use of local seed varieties and minimal mechanization.
- Crop rotation, organic fertilization, and residue management are common practices, supporting long-term soil health.
- Plot sizes are small (average ~0.03 ha), reflecting intensive farming practices aimed at self-sufficiency.

Zacapa:

- Large-scale plots dominate (average ~0.44 ha), focusing on extensive farming to maximize yield under constrained resources.
- Farmers depend heavily on fertilizers and irrigation to compensate for poor soil conditions and limited rainfall.
- Innovations like contour plowing and live barriers are increasingly adopted to combat soil erosion and improve productivity.

4. Methodology

This study aims to analyze maize growth in two contrasting regions of Guatemala—Totonicapán and Zacapa—using UAV imagery and NDVI-based metrics. The methodology is designed to capture high-resolution spatial and temporal data, process it through machine learning techniques, and integrate agronomic observations to derive actionable insights for sustainable agricultural practices. The following sections outline the key methodological steps.

"This document was reviewed using ChatGPT (OpenAI, 2025) to check text clarity and retrieve some specific references."

4.1 Study Area

Two regions were selected for their contrasting environmental and agronomic conditions:

- **Totonicapán:** Highland region with a temperate climate, high rainfall, and fertile Andisols, supporting small-scale, traditional maize farming.
- **Zacapa:** Lowland region characterized by a semi-arid climate, lower rainfall, and inceptisols, necessitating intensive resource management for maize cultivation.

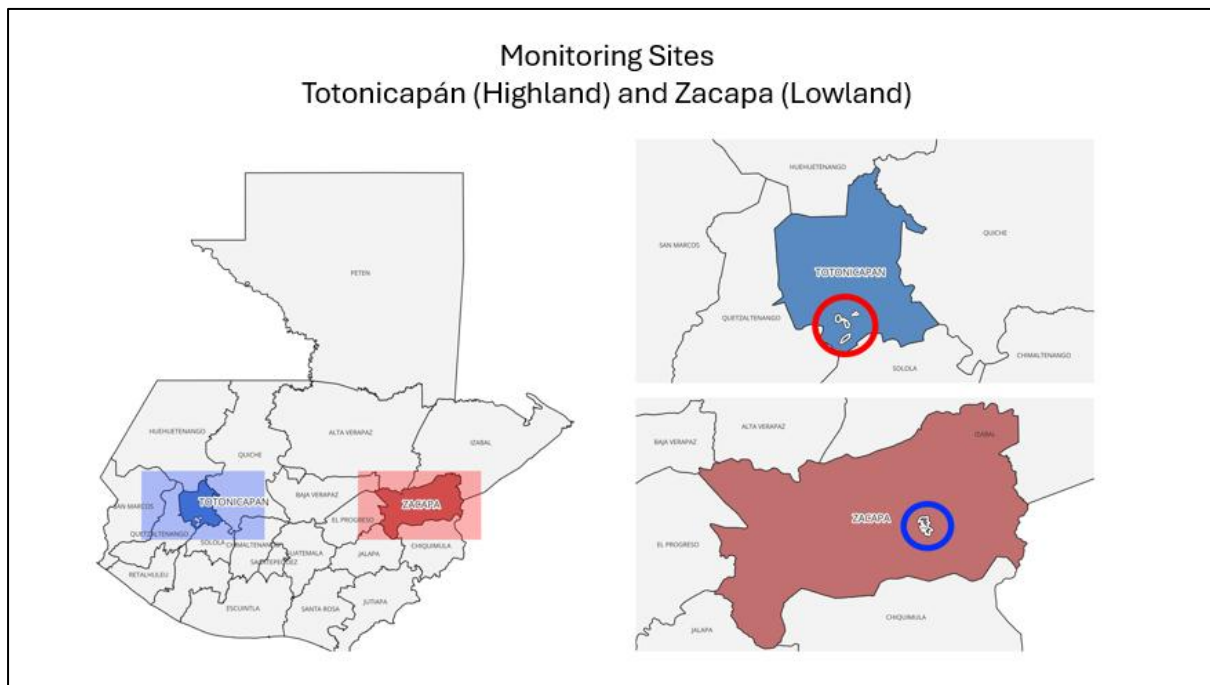


Figure 15 Study area in Guatemala (Martínez, Chávez-Can, & Navarro, 2024)

4.2 Data Acquisition

UAV Imagery:

Platform: UAVs equipped with multispectral sensors capable of capturing data in Green, Blue and NIR bands. Images were captured by Aerobots SA (Aerobot, 2025), as an external service. From the images taken, NDVI was computed using the NIR and Blue bands from UAV images.



Figure 16 Image capture preparation (Alliance Bioversity International & CIAT, 2024)



Figure 17 Evaluation corn fields Highlands (Alliance Bioversity International & CIAT, 2024)



Figure 18 Evaluation corn fields in Lowlands (Alliance Bioversity International & CIAT, 2024)

- **Vegetation Index:** Generation of NDVI using Blue and NIR bands, the following formula was use, this for fur a better characterization of crops, understanding the NDVI reflects the vegetation health, in this case, crops.

$$BNDVI = \frac{NIR - Blue}{NIR + Blue} \quad (1) \quad (\text{Gitelson \& Merzlyak, 1994})$$

- **Spatial Resolution:** 30 centimeters resolution, to determine fine-scale variability in crop fields.
- **Temporal Frequency:** Flights were conducted during two critical growth stages: the vegetative and reproductive phases, the crop cycle in highland lasts two month more than crop cycle in lowland, therefore the images were taken in different dates, thus

ensuring the same moment of growth in both locations. This coverage was done during the cultivation cycle from 2024.

Activities	May	June	July	August	September	October	November	December
Growth cycle of corn (Highland)	Planting	Germination	Growth phase I	Growth phase II	Fruiting	Harvesting		
Image capture (Highland)								
Growth cycle of corn (Lowland)	Planting	Germination	Growth phase I	Growth phase II	Fruiting			
Image capture (Lowland)								

Planting

Germination

Growth phase I

Growth phase II

Fruiting

Harvesting

Figure 19 Image capture schedule (Alliance Bioversity International & CIAT, 2024)

- **Coverage:** Selected fields in highland and lowland, covering a total area of approximately 6.5 hectares per region. The Area of Interest (AoI) was selected according to the region with more crop fields and with more farmers sharing agronomic data.
- After filtering the sample points, 40 points were used in highland (80 counting both phases) and 48 in lowland (96 in both phases). Due to weather conditions (rainy season, presence of clouds) some points were not included on the analysis, mostly in the highlands.
- Agronomic data, including yield measurements, were collected from selected plots in both region. Not all points were considered during the analysis due to atmospheric issues, clouds were making noise due to rain season at the moment when images were taken.

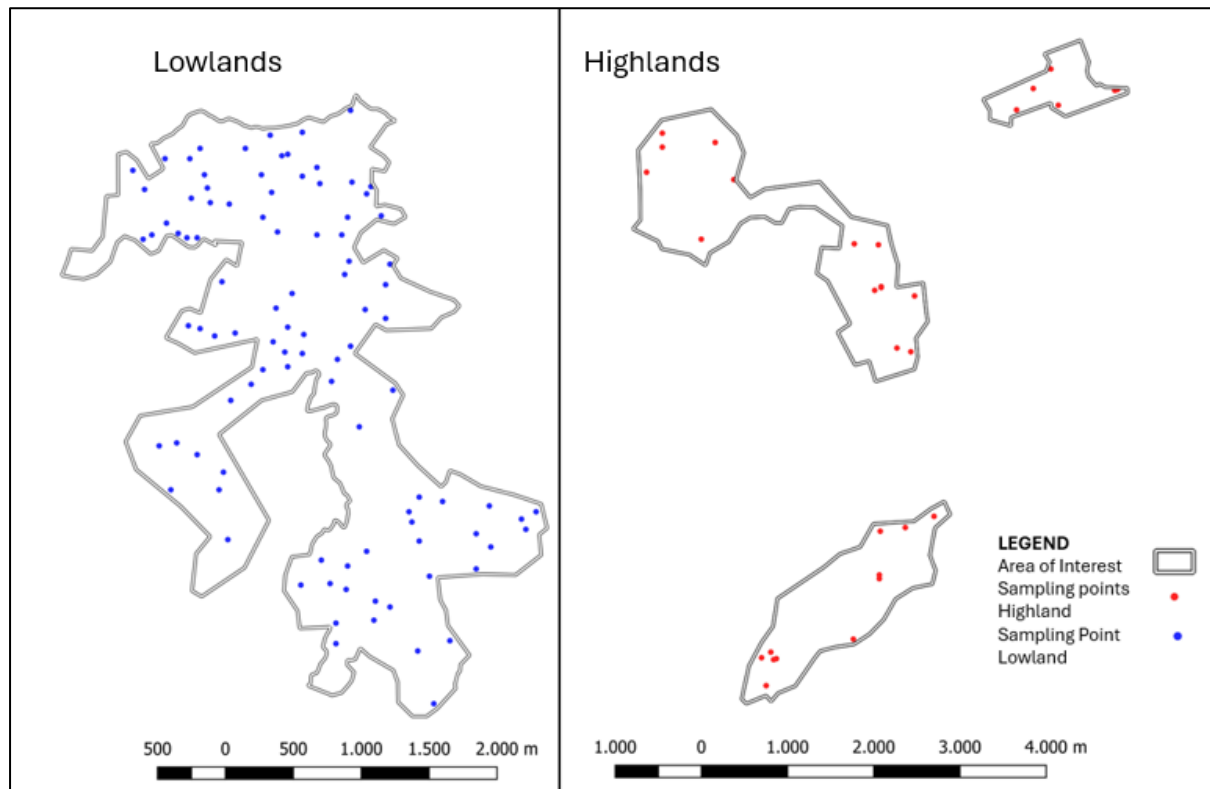


Figure 20 Sampling sites in each location (Martínez, Chávez-Can, & Navarro, 2024)

Software and tools:

- UAV Operation: Drone type was AFP 300, design and created by Aerobots. DJI Terra for flight planning and image acquisition.
- Data processing: QGIS was used for images revision, checking all raster were correctly aligned, with the necessary bands from the provider and checked with other web maps like google earth maps and open street maps.

4.3 Image Preprocessing

1. Tile Creation:

U-Net requires image tiles for processing images, handling large images directly is computationally expensive and memory-intensive. Tiling allows the model to focus on smaller regions (in this case, corn fields), improving segmentation accuracy while preserving spatial details. Considering this, the UAV imagery was divided into 64×64 pixel tiles in .tif format, overlapping the sampling polygon for efficient processing, where the tiles overlap the polygon

created from the sampling points. The NDVI band was re-scaled to **[0, 255]**, matching the structure of the other three bands. Each **.tif** tile was then converted into a **.png** image containing four-band information.


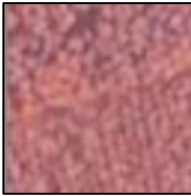
	
Tile overlapping corn field polygon	PNG image from tiff tile

Table 4 Description of tiles generation

- According with Persello & Bruzzone, 2010, If spatial analysis relies solely on vegetation indices, like NDVI, the results may lack the accuracy achieved when incorporating all original spectral bands. Each tile included four bands: Green, Red, NIR, and Blue-NDVI.

2. Labelling and Mask Creation:

Labeling is essential for training machine learning models, as it provides the ground truth needed for accurate segmentation. **Label Studio** was chosen for its flexibility and support for **semantic segmentation with polygons** (used in the labelling procedure inside label studio website) (labelstudio, 2024), making it ideal for annotating crop fields. The labeling setup was configured to segment “corn_fields”, and once completed, a **.json** file containing annotations for each tile was downloaded. **Binary masks** were then generated, where **1 represents corn fields** and **0 represents non-crop areas**, ensuring clear differentiation for model training.

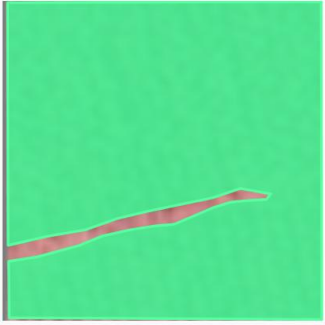
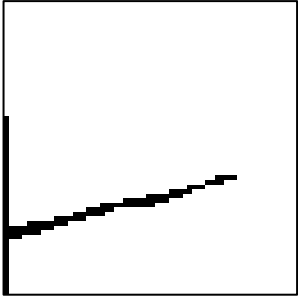
	
Label in png image.	Binary mask generated using json file.

Table 5 Description of mask generation

A python (libraries in the script: json, numpy and pycocotools) was used to convert from the json and png tiles to png binary images.

4.4 Machine Learning Model

Model Architecture

A U-Net deep learning model was employed for semantic segmentation tasks in this study. The architecture's encoder-decoder structure, equipped with skip connections, ensured the preservation of spatial details, which is critical for delineating crop fields from UAV imagery. The model was specifically optimized for agricultural imagery, leveraging the multi-band inputs (Green, Red, NIR, and Blue-NDVI).

Training, Test, and Validation

1. Data Splitting:

The dataset was divided into three subsets: training (70%), validation (15%), and testing (15%), the same data split was replicated across all growth phases and locations to ensure consistency in model training and performance assessment. This division ensures that the model learns from the majority of the data while maintaining independent datasets for hyperparameter tuning and final evaluation, preventing overfitting.

Place	Phase	Training	Testing	Validation
Highland	1	70	15	15
Highland	2	70	15	15

Lowland	1	70	15	15
Lowland	2	70	15	15

Table 6 Dataset distribution for U-Net analysis

2. Training:

The inputs for training corresponded to 70% of the dataset, consisting of 64×64 pixel TIFF tiles with 4 bands and their corresponding PNG binary masks. The training process involved feeding these labeled images into the U-Net model. Two key modules facilitated this process: “Dataloader.py”, responsible for loading, processing, and batching the dataset, and “UNet.py”, which defined the model architecture, including convolutional, pooling, and up-sampling layers.

The hyperparameters to adjust in the training section were:

Hyperparameter	Value
<i>Learning Rate</i>	Controls how much the model updates during training.
<i>Batches</i>	A subset of data processed at once for efficiency.
<i>Epochs</i>	One complete pass through the entire dataset.

Table 7 Hyperparameters in training loop, U-Net (TensorFlow, 2025) (PyTorch, 2025)

A “.pth” file, generated in this stage as an output, a different file for every processing data group (location and growth stage), the file contain all the information after processing with the training model, that file will be used in the testing and validation section

3. Testing:

The testing phase utilized 15% of the dataset, consisting of multispectral TIFF images with four bands (NIR, Green, Blue, NDVI) and their corresponding PNG binary masks. The pretrained U-Net model (saved as a “.pth” file) was used for segmentation and prediction. Each test image was processed through the U-Net model, generating a probability mask, which was then thresholded to create binary segmentation masks. The predictions were saved as PNG images, maintaining filenames matching the original input images. Finally, the predicted masks

were visually compared with the ground truth masks to assess segmentation accuracy. The output consisted of binary PNG masks highlighting crop field areas.

4. Validation:

The validation phase used the remain 15% of the dataset, consisting of multispectral TIFF images with four bands (NIR, Green, Blue, NDVI) and their corresponding PNG binary masks like all stages before. The pretrained U-Net model (".pth" file) generated during training was used for evaluation. Predictions were generated in batches, applying a sigmoid activation function to produce probability masks, which were then thresholded at 0.5 to create binary segmentation masks. A confusion matrix was computed and reorganized for interpretation, and key performance metrics—including **accuracy**, **precision**, **recall**, **F1 score**, and **Cohen's Kappa**—were calculated using flattened predictions and ground truth masks, all metrics are described in Table 7. The final metrics and confusion matrix were saved in a CSV file for further analysis.

Metric	Definition	Formula
<i>Accuracy</i>	Measures the proportion of correctly classified instances (both positive and negative) relative to the total number of instances in the dataset.	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$
<i>Precision</i>	Evaluates the fraction of instances predicted as positive that are truly positive.	$Prec = \frac{TP}{TP + FP}$
<i>Recall</i>	Recall measures the fraction of actual positive instances that were correctly identified.	$Rc = \frac{TP}{TP + FN}$
<i>F1 Score</i>	Harmonic mean of precision and recall, providing a single metric that balances the trade-offs between the two.	$F1S = 2 * \frac{Prec * Rc}{Prec + Rc}$

<i>Cohen Kappa Score</i>	Evaluates the agreement between two annotators or a model and ground truth, accounting for chance agreement.	—
--------------------------	--	---

Table 8 Metrics for validation assessment, U-Net (Sokolova & Lapalme, 2009)

5. Modules

The **Data Loader Module** (dataloader.py) manages data loading and preprocessing for training, testing, and validation stages in the U-Net model, ensuring that input images and masks are properly prepared. The module initializes by taking directories for images and masks as inputs, verifying that the number of images matches the number of masks, and ensuring filename correspondence. During data preprocessing, it loads 4-band TIFF images (NIR, Green, Blue, NDVI), normalizes all bands to the range [0, 1], and converts masks into binary tensors with values of 0 or 1 for segmentation tasks. The data is then loaded using PyTorch's Dataset and DataLoader, providing batches of image-mask pairs for efficient model training and evaluation. The inputs for this module include directory paths for images and masks, along with optional transformations such as resizing and augmentation. The final output consists of batches of image-mask pairs ready for model processing.

The **U-Net Architecture Module** (UNet.py) defines the deep learning model for semantic segmentation of crop fields using multispectral images. The model takes as input images with four spectral bands (NIR, Green, Blue, NDVI) and processes them through a structured encoder-decoder architecture. The encoder extracts important spatial features using convolutional layers and downsampling with max pooling, while the decoder reconstructs the spatial resolution using transposed convolutions and skip connections to merge encoder outputs. The input images pass through the encoder layers to capture hierarchical features, and the decoder reconstructs the segmentation masks, producing delineated crop fields in the final output. The output of the module consists of predicted segmentation masks for each input image, allowing for accurate crop field segmentation and classification.

6. Software and tools

All the U-Net three sections were performed using python (libraries: torch, torch.utils.data, torchvision.transform and sklearn.metrics) in Visual Studio.

4.5 NDVI Analysis

The **NDVI analysis** involved extracting key metrics, including **mean, maximum, minimum, and standard deviation** from each tile, however before this analysis, a Shapiro-Wilk test and t-test were done, this for ensuring normal distribution and statistical difference in between each in each region

To evaluate regional differences, an independent t-test was conducted to compare NDVI metrics between Totonicapán (highland) and Zacapa (lowland) across both growth stages. This statistical test determined whether observed differences in NDVI were due to random variation or actual disparities in crop growth conditions influenced by environmental factors. The data structure included NDVI mean values extracted from UAV imagery, organized into two groups—highland NDVI and lowland NDVI, as separated dataset—for statistical comparison. Additionally, the **Shapiro-Wilk** test was applied to assess whether the NDVI dataset followed a normal distribution, ensuring that statistical conclusions were accurate and avoiding misinterpretations. The normality test was performed separately for each region to validate the dataset before proceeding with further statistical analyses.

Variable	Formula	Description
○ Mean	$\bar{X} = \frac{\sum X_i}{n}$	Average of all NDVI values.
○ Variance	$s^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$	Spread of NDVI Values.
○ Weights	$a_i = \frac{m}{\sqrt{\sum m^2}}$	Normalized order statistic (m represents expected values for a normal distribution).
○ W-Statistics	$W = \frac{(\sum a_i * X_{(i)})^2}{\sum (X_i - \bar{X})^2}$	Measures deviation from normality.
○ p-Value	Derived from Shapiro-Wilk test	Indicates statistical significance.

Table 9 Shapiro - Wilk test variables (Royston, 1992)

After computation we have to evaluate the results by rejecting or accepting a hypothesis, and make a decision, to know if the data is normal distributed.

<u>Hypotheses</u>	<u>Decision Rule</u>
-H ₀ (Null Hypothesis): The data follows a normal distribution.	If p-value > 0.05 , we fail to reject H₀ Data is normally distributed.
-H _A (Alternative Hypothesis): The data does not follow a normal distribution.	If p-value < 0.05 , we reject H₀ Data significantly deviates from normality.

The **t-test** (Student's t-test) is used to **compare the means** of two independent groups (NDVI in highland vs. lowland) and determine if the difference is statistically significant. This is done by following the procedure in Table 9.

Variable	Formula	Description
Mean	$\bar{X} = \frac{\sum X_i}{n}$	Average of NDVI values per location.
Variance	$s^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$	Spread of NDVI values.
Pooled Variance	$s_p^2 = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}$	Combined variance of both samples.
Standard Error	$SE = \sqrt{s_p^2 * (\frac{1}{n_1} + \frac{1}{n_2})}$	Measure variability between means.
T-Statistic (t-score)	$t = \frac{\bar{X}_1 - \bar{X}_2}{SE}$	Differences in means relative to SE.
Degrees of Freedom	$df = n_1 + n_2 - 2$	Determine critical t-value lookup
p-Value	Derived from t-distribution	Probability of observing <i>t</i> .

Table 10 T-test variables (Montgomery, 2013)

Once we have computed all necessary variables, we must evaluate the results by rejecting or accepting a hypothesis, by making a decision, to know if there is a significant statistical difference between mean NDVI values from both locations.

<u>Hypotheses</u>	<u>Decision Rule</u>
-H ₀ (Null Hypothesis): The mean NDVI for highland and lowland are equal.	If p-value < 0.05 , we reject H₀ The difference in NDVI is statistically significant.
-H _A (Alternative Hypothesis): The mean NDVI for highland and lowland are different.	If p-value > 0.05 , we fail to reject H₀ No significant difference between NDVI means.

Once Shapiro-Wilk test and t-test were performed, a statistical dispersion by place chart have to be done for a proper description using NDVI statistical values. Temporal NDVI trends were analyzed to track phenological transitions, with stage 1 and stage 2, all NDVI metrics (mean, min, max and stdev) will demonstrate a specific behavior for each place and each stage analyzed.

The extraction and t-test process were performed using python (Stats extraction libraries: geopandas, rasterio, t-test libraries: pandas, numpy, scipy.stats) in Visual Studio.

4.6 Agronomic Data

Yield Integration Analysis:

NDVI metrics were related with yield data to evaluate the relationship between vegetation vigor and productivity, this was done to complete the site description by including agronomic data. NDVI could serve as an indicator of maize performance across different growth stages, agronomics activities and environmental conditions.

5 Results

This section highlights the performance of the U-Net model in segmenting crop fields in two contrasting environment and two corn growing stages, the analysis of NDVI metrics, and the integration of agronomic data to evaluate regional disparities between Totonicapán (highland) and Zacapa (lowland).

5.1 Performance of the U-Net Model, image segmentation

5.1.1 Training

The training section demonstrates the effectiveness of the model capturing cornfields from the images. In the highland region, both growth stages were trained over 100 epochs, achieving high values for metrics such as Accuracy, Dice Coefficient, and Intersection over Union (IoU), alongside low Loss values. These results indicate that the model has successfully learned to differentiate cornfield patterns with minimal errors in this section. The high performance can be attributed to the distinct spectral and structural contrast between corn crops and other vegetation types in the highland region.

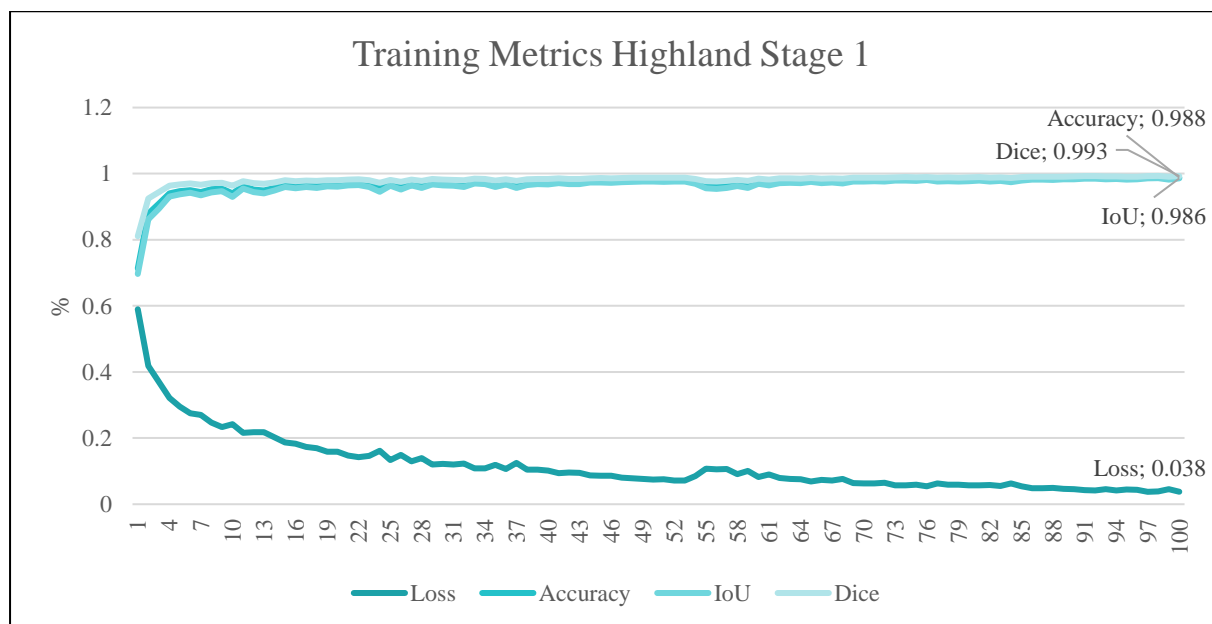


Figure 21 Training metrics for evaluation, Highlands Stage 1

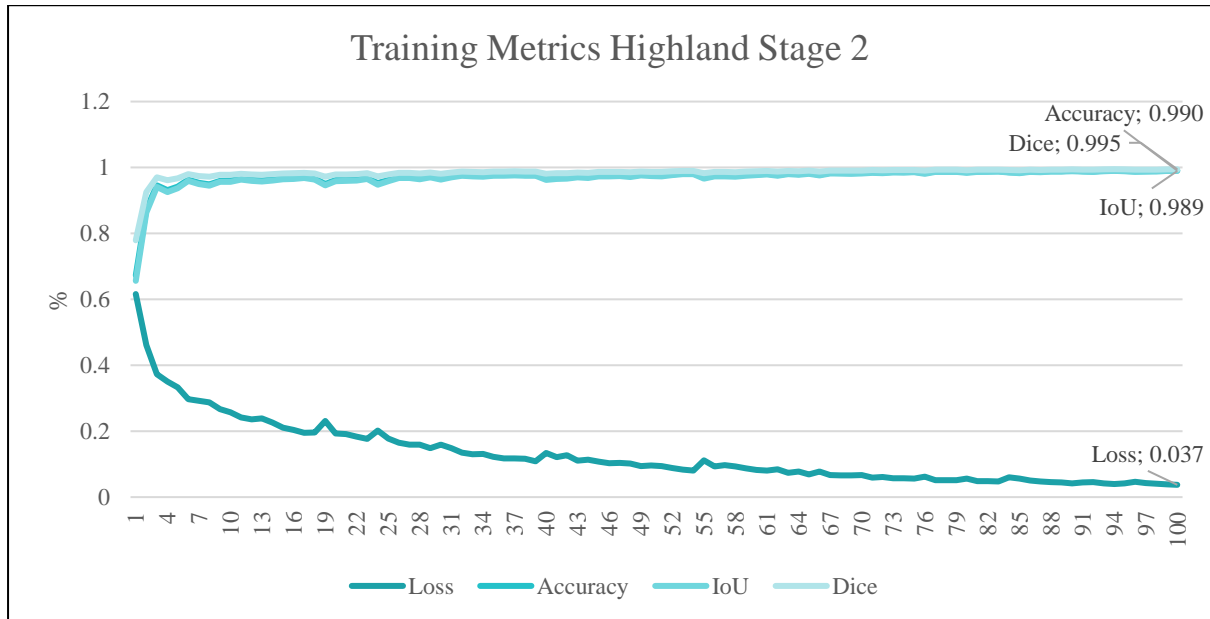


Figure 22 Training metrics for evaluation, Highlands Stage 2

On the other hand, in the lowland region, the final values for Accuracy, Dice Coefficient, IoU, and Loss after 100 epochs are also good. However, during the training process, noticeable fluctuations were observed in these metrics (Figure 21 & Figure 22). These jumps suggest that the model faced challenges distinguishing cornfields from other types of vegetation. This difficulty could be attributed to factors such as lower contrast between crops and surrounding vegetation, higher variability in spectral signatures, or increased noise in the lowland imagery. Despite these challenges, the model achieved reasonable performance in the lowland environment.

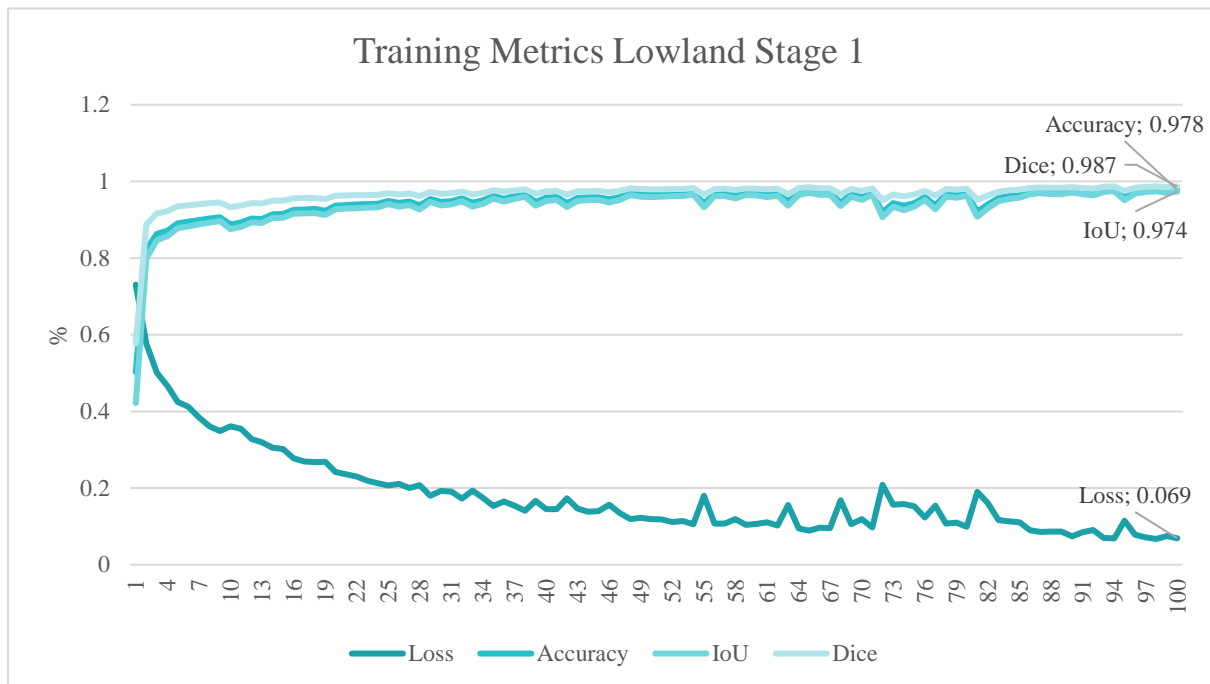


Figure 23 Training metrics for evaluation, Lowlands Stage 1

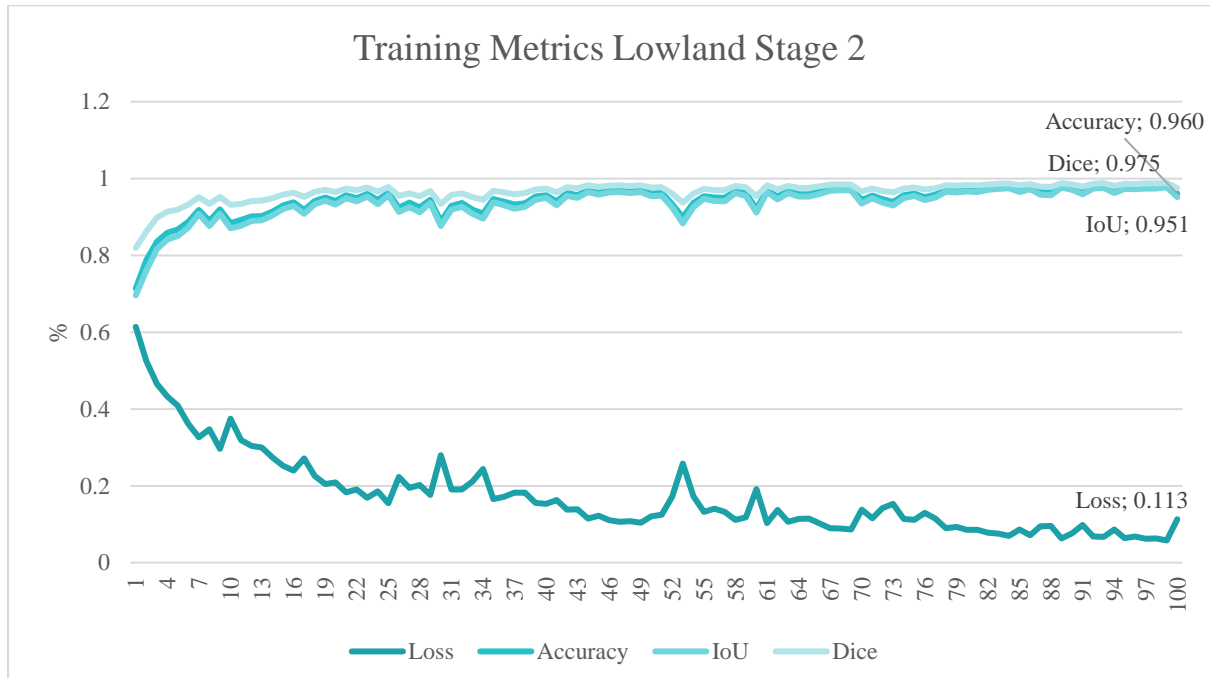


Figure 24 Training metrics for evaluation, Lowlands Stage 2

5.1.2 Validation

The validation metrics were derived from the confusion matrices generated for each location and growth stage. From this matrices, five different metrics were computed in order to understand the performance of the model in each condition (place and growth stage).

Highland 1			Highland 2		
	Predicted positive	Predicted negative		Predicted positive	Predicted negative
Actual positive	21054	1789	Actual positive	21941	868
Actual negative	1042	691	Actual negative	606	1161

Lowland 1			Lowland 2		
	Predicted positive	Predicted negative		Predicted positive	Predicted negative
Actual positive	12379	3100	Actual positive	12120	4226
Actual negative	2169	2832	Actual negative	1501	2633

Figure 25 Confusion matrix for each location and each stage

Highlands:

The model's performance in the highlands is notably strong, particularly in stage 2 (HL2). Metrics for HL2 show **Accuracy (0.94)**, **Precision (0.96)**, **Recall (0.97)**, **F1 Score (0.97)**, and **Kappa Coefficient (0.58)**. These results combined demonstrate the effectiveness of the model identify cornfields with minimal misclassifications. Below is a detailed explanation of how each metric reflects specific aspects of the model's behavior:

- **Accuracy:** This high value indicates that 94% of all predictions (both positive and negative) match the ground truth. It reflects the overall reliability of the model in segmenting cornfields (rare errors from the model). The distinct spectral characteristics in the highlands help the model make clear distinctions between corn crops and other land features (other type of vegetation).
- **Precision:** The high value here reflects the model's ability to minimize false positives, likely due to the well-defined boundaries and unique reflectance of corn crops in the highlands. This behavior indicates that the model is less likely to misclassify other vegetation as cornfields.

- **Recall:** Sensitivity of the model, showing how effectively it captures all true positives. A recall of 0.97 suggests that the model identifies nearly all cornfield pixels correctly, which could be due to the high visibility of mature crops and the distinct growth patterns in stage 2.
- **F1 Score:** Balance between precision and recall, confirming that the model is improving in both avoiding false positives and capturing true positives. This metric highlights the robustness of the model in highland.
- **Kappa Coefficient:** Kappa is lower than other metrics, its improvement from stage 1 (0.27) suggests that stage 2's distinct crop patterns enhance the model's reliability in capturing complex pixel relationships.

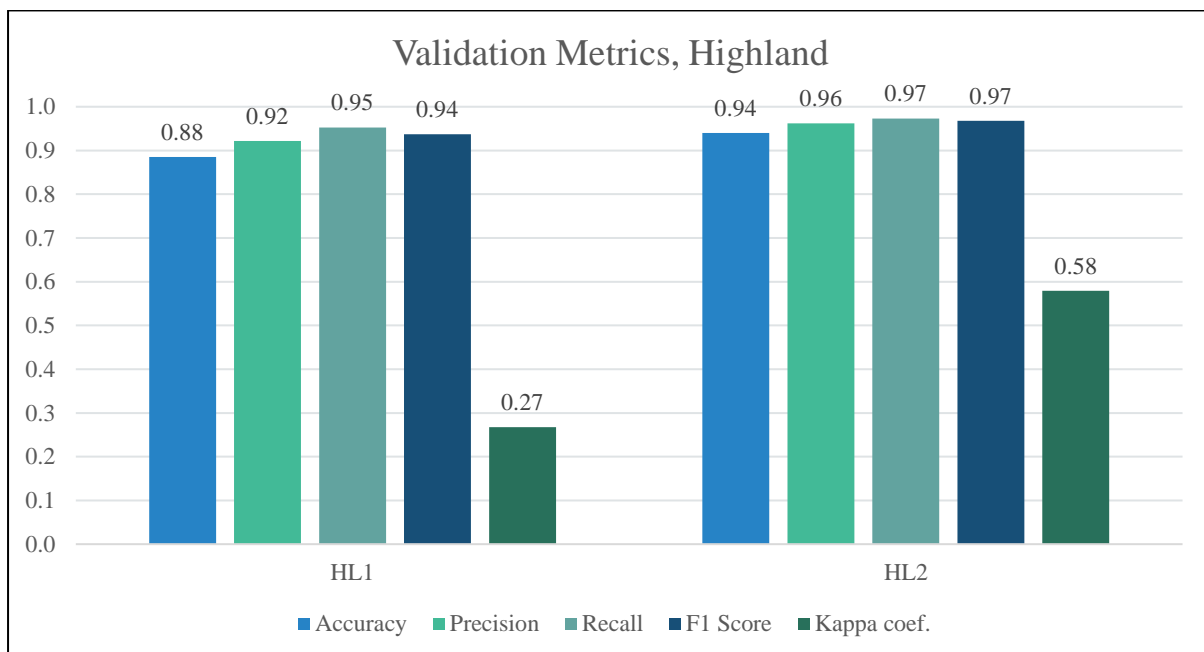


Figure 26 Validation metrics, Highlands

The improvement from stage 1 to stage 2 indicates that crop maturity in stage 2 creates more distinct features, such as larger leaf canopies and uniform growth stages, which are easier for the model to segment (Figure 11).









Labelled	 Totonicapan_f2_merged_18_tile_1 28_test02_mask.png	 Totonicapan_f2_merged_34_tile_1 28_test02_mask.png	 Totonicapan_f2_merged_38_tile_1 28_test02_mask.png	 Totonicapan_f2_merged_39_tile_1 28_test02_mask.png
Predicted	 Totonicapan_f2_merged_18_tile_1 28_test02_pred_mask.png	 Totonicapan_f2_merged_34_tile_1 28_test02_pred_mask.png	 Totonicapan_f2_merged_38_tile_1 28_test02_pred_mask.png	 Totonicapan_f2_merged_39_tile_1 28_test02_pred_mask.png

Table 11 Predictions vs Ground Truth, Highlands

Lowlands:

In the lowlands, the performance of the model is lower in comparison with the Highland, in particular analyzing stage 2 (LL2). Metrics for LL2 reveal Accuracy (0.72), Precision (0.74), Recall (0.89), F1 Score (0.81), and Kappa Coefficient (0.30). These results highlight specific challenges in this environment and reveal how each metric reflects the model's struggles:

- **Accuracy:** The lower accuracy indicates that only 72% of predictions align with the ground truth. This reflects a higher error rate compared to the highlands and suggests that the lowland environment presents more complex patterns, such as mixed vegetation or inconsistent growth stages.
- **Precision:** Precision is lower in LL2, indicating a higher occurrence of false positives. This could be caused by overlapping spectral signatures between cornfields and surrounding vegetation (weeds or similar crop types), which confuse the model.
- **Recall:** Despite the challenges in previous metrics, the recall is relatively high, showing that the model captures most true positives. However, this comes at the expense of precision, as the model may classify non-cornfield pixels as cornfields to avoid missing true positives. Focusing on sensitivity rather than specificity.

- **F1 Score:** The moderate value highlights that the model slightly struggles to maintain both high sensitivity and specificity in this environment. This indicates that the lowland conditions, such as variability in crop appearance, challenges the model's ability.
- **Kappa Coefficient (0.30):** The low Kappa value signals weak agreement between the predictions and ground truth beyond chance. This weak agreement underscores the difficulty the model faces in identifying consistent patterns in the lowlands, likely due to greater environmental variability and noise in the imagery.

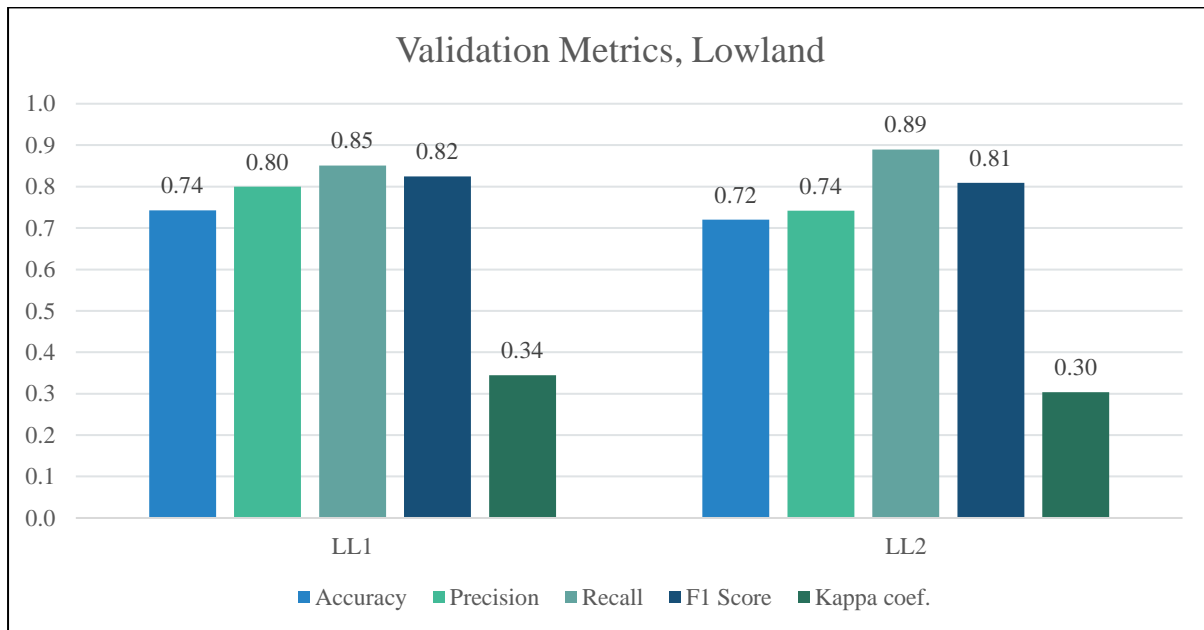


Figure 27 Validation metrics, Lowlands stage 1 and 2

The performance in stage 1 (LL1: Accuracy 0.74, Precision 0.80, Recall 0.85, F1 Score 0.82, Kappa 0.34) is slightly better but still reflects significant variability. The lower precision and Kappa values across both stages suggest that lowland crops lack the clear spectral and structural distinctions found in the highlands, making them harder for the model to classify accurately. Even with the metrics in Lowlands, the values were good enough for the extraction of cornfields from images, allowing to continue with the characterization.









Labelled	 Zacapa_f2_merged_21_tile_128_test02_mask.png	 Zacapa_f2_merged_31_tile_128_test02_mask.png	 Zacapa_f2_merged_43_tile_128_test02_mask.png	 Zacapa_f2_merged_48_tile_128_test02_mask.png
Predicted	 Zacapa_f2_merged_21_tile_128_test02_pred_mask.png	 Zacapa_f2_merged_31_tile_128_test02_pred_mask.png	 Zacapa_f2_merged_43_tile_128_test02_pred_mask.png	 Zacapa_f2_merged_48_tile_128_test02_pred_mask.png

Table 12 Predictions vs Ground Truth, Lowlands

5.2 Regional Comparison and T-Test Results

5.2.1 NDVI Distribution for Two Locations:

The histogram (Figure 28) shows the normalized difference vegetation index (NDVI) distributions for two locations (highland and lowland) across growth phases 1 and 2.

- **Highland NDVI** (in blue) exhibits a tighter spread and a higher mean value (0.341), suggesting healthier or denser than Lowland.
- **Lowland NDVI** (in red) is more spread out and has a lower mean (0.301), which might indicate greater variability in vegetation health and coverage.

There is significant overlap between the two distributions in the range of 0.25–0.35 NDVI. This overlap reflects the shared characteristics between the locations during these growth phases.

Vertical lines mark the mean NDVI values are marked for each location:

- **Highland Mean:** 0.341 (blue line)
- **Lowland Mean:** 0.301 (red line)

The clear separation of these means reinforces the statistical difference between the two locations, as confirmed by the T-test results. Ensuring Highland and Lowland can be treated as different groups of values.

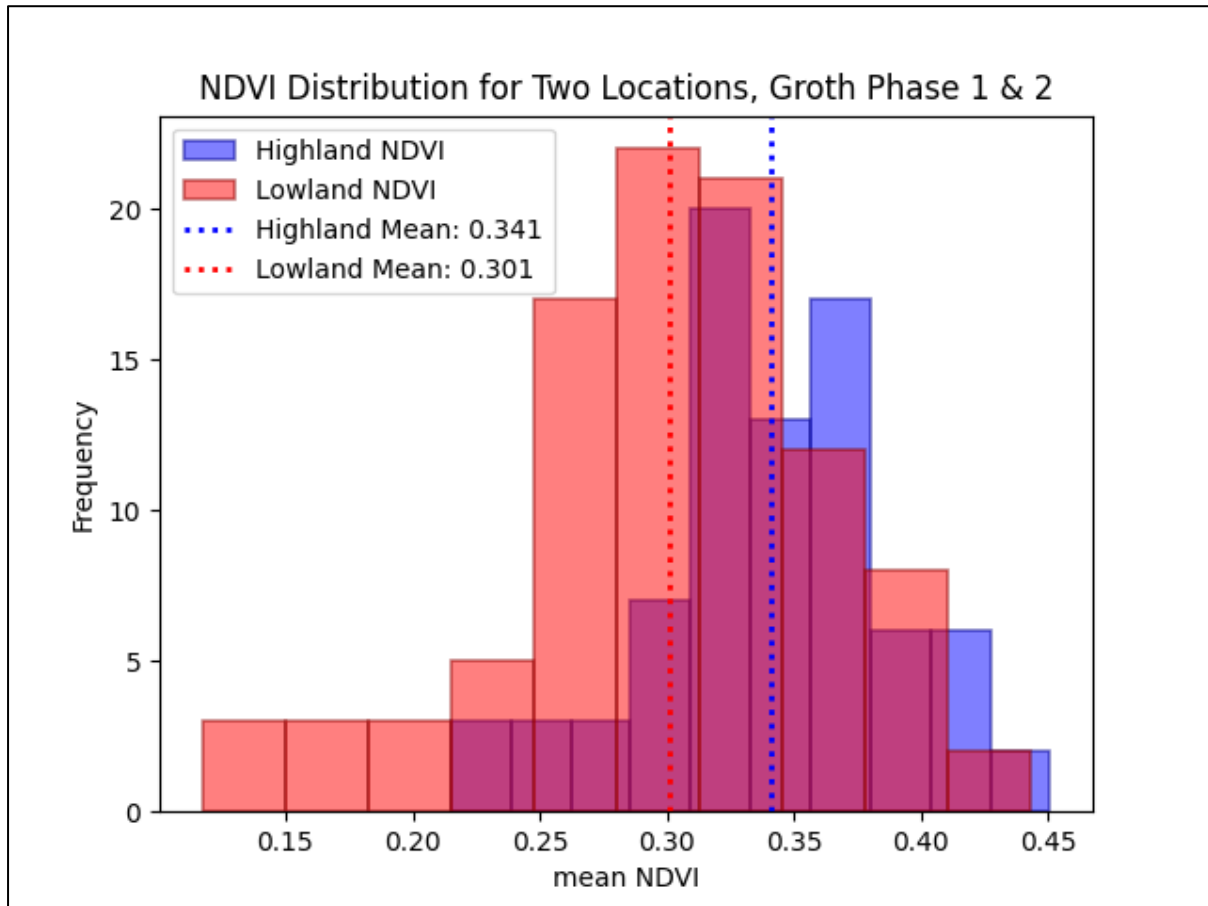


Figure 28 NDVI distribution in Highlands and Lowlands

5.2.2 Statistical Test Results

The Shapiro-Wilk test was conducted to assess whether NDVI values in the highland and lowland regions follow a normal distribution, which is a key assumption for conducting a t-test.

Highland NDVI: $W = 0.9789$, $p = 0.20$

- Since $p > 0.05$, we fail to reject the null hypothesis, indicating that NDVI values in the highland region show significant normality. The data distribution aligns with the expected normal pattern, making it suitable for parametric statistical tests.

Lowland NDVI: $W = 0.9701$, $p = 0.0272$

- The **p-value** < 0.05 suggests rejecting the null hypothesis, implying that NDVI values in the lowland region deviate slightly from a normal distribution. However, the W value (0.9701) is still close to 1, indicating a distribution that is largely normal except for minor deviations.

While the **p-value** for the lowland NDVI suggests non-normality, visual inspection of the data (Figure 28) reveals that deviations primarily occur in the lower tail of the distribution. Given that the sample size is relatively small (96 and 80), the Shapiro-Wilk test might be overly sensitive to minor irregularities, causing it to reject normality even when the distribution is approximately normal for practical purposes.

In statistical practice, small sample sizes (lower than 100 observations) can lead to false rejections of normality due to inherent variability rather than meaningful departures from the normal distribution. Since the t-test is robust to slight violations of normality, particularly when sample sizes are moderate, we can reasonably justify proceeding with the independent t-test to compare NDVI values between highlands and lowlands.

T-Test (Comparing Means):

- T-statistic=4.4823, **p-value**= $1.335 \cdot 10^{-5}$

This indicates a highly significant difference in mean NDVI values between the two regions. Since the p-value is less than 0.05, we reject the null hypothesis, confirming that the observed variation in NDVI is unlikely to be due to random chance. This result supports the finding that vegetation in highland regions exhibits higher NDVI values on average compared to lowland regions, reinforcing the hypothesis that these areas have distinct growth patterns influenced by environmental conditions. The statistical significance of this difference validates the use of NDVI as a reliable metric for distinguishing vegetation health between the two regions.

5.3 NDVI Metrics Analysis

5.3.1 Mean NDVI Trends

Totonicapán (Highland):

In Phase 1, the mean NDVI is **0.33**, which increases slightly to **0.35** in Phase 2. This gradual improvement suggests that vegetation in the highlands becomes marginally healthier or denser as the crops mature. The small standard deviation (~ 0.06) reflects a high level of consistency in vegetation health across the cornfields, with minimal variability among samples. This stability indicates uniform environmental conditions and well-defined vegetation patterns, which are characteristic of the highland region and contribute to the model's strong performance in this area.

Zacapa (Lowland):

In Phase 1, the mean NDVI is 0.26, which a significant increases to 0.34 in Phase 2. This larger increase, compared to Totonicapán, suggests a more pronounced vegetative improvement, potentially driven by better crop establishment or a reduction in environmental stress. The standard deviation (~ 0.05) is slightly smaller than that of Totonicapán, indicating consistent performance across samples. However, the results still highlight challenges in achieving higher NDVI values, likely due to the environmental variability and stressors characteristic of the lowland region.

5.3.2 Minimum and Maximum NDVI

Minimum

In highland, the minimum NDVI increases from **0.077** in Phase 1 to **0.153** in Phase 2, reflecting an improvement in baseline vegetation health. In contrast, The lowland shows a negative minimum NDVI value (**-0.002**) in Phase 1, which rises slightly to **0.060** in Phase 2. The negative value in Zacapa suggests the presence of barren areas, sparse vegetation, or potential noise in the data during Phase 1, highlighting the variability and challenges in vegetation conditions in the lowland region.

Maximum NDVI:

The values in highland remain consistently high across both phases, decreasing slightly from **0.571** in Phase 1 to **0.562** in Phase 2. This indicates consistently healthy vegetation in optimal areas. Meanwhile, in lowland, NDVI increases from **0.505** in Phase 1 to **0.605** in Phase 2, reflecting a notable improvement in vegetation health in certain areas as crops mature and conditions improve.

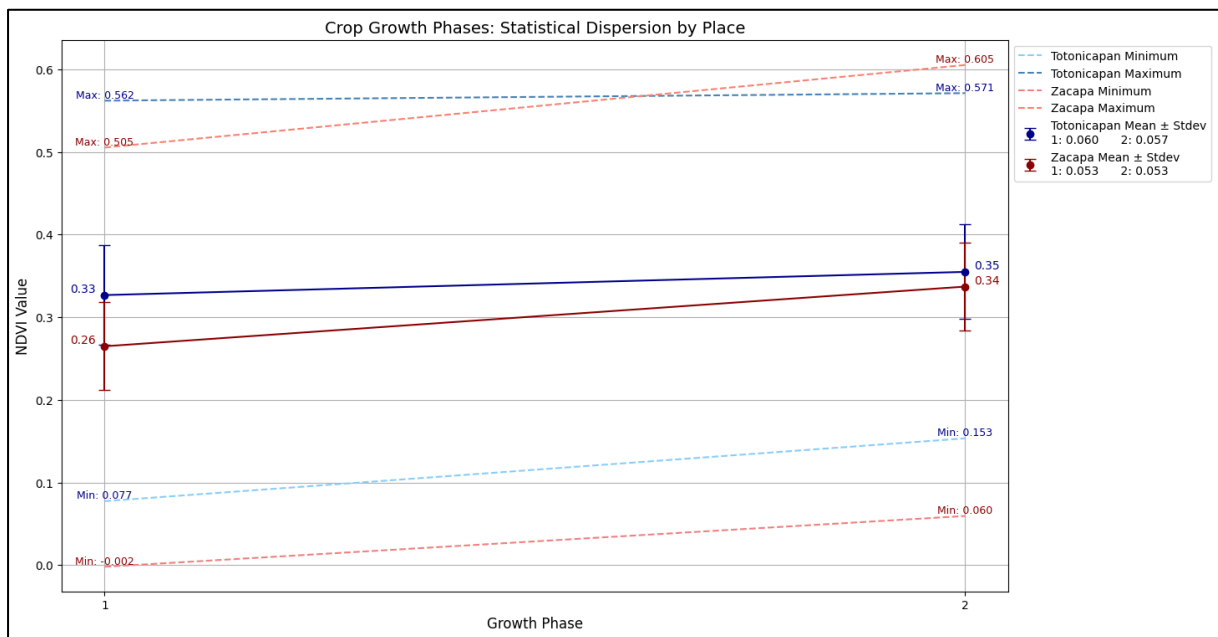


Figure 29 Crop growth stages with statistical dispersion by location

5.3.3 Dispersion and Variability

Totonicapán: The NDVI distribution shows limited dispersion (narrow error bars) and remains consistently higher, indicating stable vegetation health across the highland environment.

Zacapa: The wider dispersion and larger increase in maximum NDVI suggest greater variability in vegetation health. This could reflect a mix of areas with significant improvement and others that remain under environmental stress.

5.4 Agronomic Data Integration

For each sampled point, the corn yield per unit area was recorded, providing valuable data for assessing crop productivity. This yield information offers crucial context for interpreting NDVI behavior. In Totonicapán (highland), the yield is **0.5 t/ha**, while in Zacapa (lowland), the yield is significantly higher at **2 t/ha**. These values, though contrasting, align with expectations based on the NDVI results. For the specific NDVI values observed—**0.33 to 0.35** in Totonicapán and **0.26 to 0.34** in Zacapa—we can infer that these yields represent the average productivity across all samples within each region. The higher yield in Zacapa, despite lower NDVI values, suggests that environmental factors, farming practices, or input levels in the lowlands contribute significantly to productivity. Conversely, the lower yield in Totonicapán, despite healthier vegetation indicated by higher NDVI, may result from limitations such as lower temperatures, less precipitation during the crop cycle. This integration of agronomic and NDVI data highlights the intricate relationship between vegetation health, environmental conditions, and crop yield. It emphasizes the value of remote sensing tools and data analysis in providing insights for agronomic practices.

6. Conclusions

This study successfully demonstrated the application of U-Net-based image segmentation in delineating cornfields in two contrasting environments, reinforcing the capabilities of deep learning models in UAV-based crop monitoring. Segmentation accuracy was significantly higher in highland regions, where clear spectral contrast improved classification, while lowland segmentation faced greater challenges due to complex vegetation conditions and spectral variability. Despite these limitations, U-Net successfully extracted NDVI values, proving its robustness as a precision agriculture tool for assessing crop health.

Beyond segmentation, the **extraction and statistical analysis of NDVI metrics** provided deeper insights into crop vigor and growth differences across the two regions. The results showed that NDVI values were consistently higher in the highlands, while the lowland NDVI distribution displayed greater variability, reflecting environmental stressors. The application of a t-test confirmed that these differences were statistically significant, supporting the hypothesis that cornfields in these contrasting regions exhibit distinct growth behavior. However, yield analysis revealed an unexpected trend, where higher NDVI did not directly correspond to higher productivity, emphasizing the need to consider additional agronomic and environmental factors when interpreting remote sensing data.

Future research should focus on enhancing segmentation accuracy in spectrally complex environments by integrating hybrid deep learning models. Incorporating the red band is essential for a more detailed analysis of maize growth stages. A comparative experiment using U-Net with and without the red band should be conducted to evaluate its impact on accuracy and precision. Additionally, to better characterize maize development over time, more UAV flights are required, with a minimum of five acquisitions, one for each key growth stage, to capture temporal variations in crop vigor and phenology.

7. References

- Aerobot. (2025, 01 23). *www.aerobots.gt*. Retrieved from *www.aerobots.gt*:
<https://www.aerobots.gt/en/home/soluciones/recoleccion-y-analisis-de-datos/>
- Alliance Bioversity International & CIAT. (2024, 06). Monitoreo de cultivos con imágenes de drones. Ciudad de Guatemala, Guatemala: CIAT.
- Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). *Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation*. IEEE.
- Alvino, F. C., Alemán, C. C., Filgueiras, R., Althoff, D., & Cunha, F. F. (2020). *VEGETATION INDICES FOR IRRIGATED CORN MONITORING*. Minas Gerais: EAgrí.
- Bakken, M., Ponnambalam, V. R., From, P. J., & Moore, R. J. (2021). *Robot-supervised Learning of Crop Row Segmentation*. ResearchGate.
- Canata, T. F., Wei, M. C., Maldaner, L. F., & Molin, J. P. (2021). *Sugarcane Yield Mapping Using High-Resolution Imagery Data and Machine Learning Technique*. Sao Paulo: remote sensing.
- Cheng, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). *DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs*. IEEE.
- CIAT. (2023). *Agricultural Resilience in Latin America and the Caribbean*. Cali: CIAT.
- Cicek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation*. Freiburg.
- Copacenaru, O., Stoica, A., Catucci, A., Vendictis, L. D., Tricomi, A., Rogotis, S., & Marianos, N. (2021). *Copernicus Data and CAP Subsidies Control*. ResearchGate.
- Fisher, R. A. (1925). *Statistical Methods for Research Coworkers*. Edinburg: Oliver and Boyd.

- Gitelson, A., & Merzlyak, M. N. (1994). *Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus hippocastanum L. and Acer platanoides L. Leaves. Spectral Features and Relation to Chlorophyll Estimation*. Moscow: J. Plant Physiology.
- He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). *Momentum Contrast for Unsupervised Visual Representation Learning*. IEEE.
- INSIVUMEH. (2023). *Estado del clima en Guatemala 2022*. Ciudad de Guatemala: INSIVUMEH.
- Jones, D. A., Smith, R. T., & Lee, C. M. (2021). *Early-Season NDVI and Soil Moisture Integration for Wheat Yield Prediction*. Iowa: Journal of Agricultural Science.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). *Deep learning in agriculture: A survey*. ELSEVIER.
- labelstudio. (2024, 12 10). *labelstud.io*. Retrieved from <https://labelstud.io/http://localhost:8080/projects/?page=1>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning*. Montreal: ResearchGate.
- Li, G., Wang, J., Wang, Y., & Qi, Z. (2024). *Integration of Remote Sensing and Machine Learning for Precision Agriculture: A Comprehensive Perspective on Applications*. Lanzhou: Agronomy.
- Liu, Y., Song, W., & Deng, X. (2016). *Changes in crop type distribution in Zhangye City of the Heihe River Basin, China*. Beijing: ELSEVIER.
- Ma, Z., Wang, G., Yao, J., Huang, D., Tan, H., Jia, H., & Zou, Z. (2023). *An Improved U-Net Model Based on Multi-Scale Input and Attention Mechanism: Application for Recognition of Chinese Cabbage and Weed*. Basel: Sustainability.
- Martínez, E., Chávez-Can, M., & Navarro, C. (2024). *Evaluación piloto de diagnóstico agronómico asistido por drones en sistemas de agricultura a pequeña escala*. Guatemala City: CIAT.
- Ministerio de Agricultura, G. y. (2019). *Mapa de Clasificación Taxonómica de Suelos de Guatemala*. Guatemala: MAGA.

- Ministerio de Agricultura, G. y. (2020). *Mapa de Cobertura Vegetal y Uso de la Tierra*. Guatemala: MAGA.
- Montgomery, D. C. (2013). *Design and Analyzis of Experiments*. Arizona: Wiley.
- Motis, T. N. (2003). *Statistical Analysis of Simple Agricultural Experiments*. ECHOcommunity.
- NATIONS, U. (2024, 01 20). *UNITED NATIONS* . Retrieved from [sdgs.un.org: https://sdgs.un.org/goals](https://sdgs.un.org/goals)
- Navarro, C. (2023, 11). Asesoramiento agrícola inclusivo habilitado digitalmente para la gestion de riesgos. Ciudad de Guatemala, Guatemala.
- Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., . . . Rueckert, D. (2018). *Attention U-Net: Learning Where to Look for the Pancreas*. MICCAI.
- OpenAI. (2025, February 11). <https://openai.com>. Retrieved from <https://openai.com: https://chatgpt.com/>
- Persello, C., & Bruzzone, L. (2010). *A Novel Protocol for Accuracy Assessment in Classification of Very High Resolution Images*. Trento: IEEE.
- PIX4D. (2025, 02 05). [pix4d.com](https://www.pix4d.com). Retrieved from <https://www.pix4d.com: https://www.pix4d.com/blog/pix4dfields-vegetation-indices-for-precision-agriculture/>
- PyTorch. (2025, 01 23). pytorch.org. Retrieved from pytorch.org: https://pytorch.org/tutorials/beginner/basics/data_tutorial.html
- Ranagalage, M., Estoque, R. C., & Murayama, Y. (2018). *An Urban Heat Island Study of the Colombo Metropolitan Area, Sri Lanka, Based on Landsat Data (1997–2017)*. International Journal of Geo-Information.
- Ronneberger, O., Fischer, P., & Brox, T. (2015, 05 18). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. Heidelberg: Springer. Retrieved from Cornell University: <https://arxiv.org/abs/1505.04597>

- Royston, P. (1992). *Approximating the Shapiro-Wilk W-test for Non-Normality*. London : SPRINGER.
- Shanahan, J., Francis, D. D., Schepers, J. S., & Varvel, G. E. (2001). *Use of Remote-Sensing Imagery to Estimate Corn Grain Yield*. Negraska: ResearchGate.
- Sokolova, M., & Lapalme, G. (2009). *A systematic analysis of performance measures for classification tasks*. Montreal: ELSEVIER.
- Sui, B., Jiang, T., Zhang, Z., & Pan, X. (2020). *ECGAN: An improved conditional generative adversarial network with edge detection to augment limited training data for the classification of remote sensing images with high spatial resolution*. IEEE.
- Sustainable Agriculture Research and Education. (2017). *How to Conduct Research on Your Farm or Ranch*. SARE.
- Tao, H., Feng, H., Xu, L., Miao, M., Long, H., Yue, J., . . . Fan, L. (2020). *Estimation of Crop Growth Parameters Using UAV-Based Hyperspectral Remote Sensing Data*. Beijing: Sensors.
- TensorFlow. (2025, 01 23). <https://www.tensorflow.org/>. Retrieved from https://www.tensorflow.org/addons/tutorials/optimizers_cyclicallearningrate?hl=es-419
- Tucker, C. J. (1978). *Red and photographic infrared linear combinations for monitoring vegetation*. Meryland: ELSEVIER.
- Viana, C. M., Freire, D., Santos, J. M., & Abrantes, P. (2021). *Evaluation of the factors explaining the use of agricultural land: A machine learning and model-agnostic approach*. Lisboa: ELSEVIER.
- Wang, F., Zhang, M., Wang, X., Ma, X., & Lui, J. (2020). *Deep Learning for Edge Computing Application: A State of The Art Suvey*. IEEE.
- World Food Program. (2025, 02 13). *WFP (World Food Program)*. Retrieved from Guatemala: How WFP-supported farmers keep school learners nourished:

<https://www.wfp.org/stories/guatemala-how-wfp-supported-farmers-keep-school-learners-nourished>

- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). *SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers*. Hong Kong: NeurIPS.
- Zhang, C., & Kovacs, J. M. (2012). *The application of small UAVs in precision agriculture: A review*. Ottawa: Springer.
- Zhang, J., He, Y., Yuan, L., Liu, P., Zhou, X., & YanboHuang. (2019). *Machine Learning-Based Spectral Library for Crop*. Hangzhou: Agronomy.
- Zhang, Z., & Liu, Q. (2018). *Road extraction by deep residual U-Net*. Beihang: IEEE.
- Zhao, C., Yang, G., Liu, J., Li, Z., & Huang, Y. (2017). *Unmanned aerial vehicle remote sensing for field-based crop phenotyping: Current status and perspectives*. Lausanne, Switzerland.: Frontiers Media SA.
- Zhou, Z., Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*. Arizona: Springer Nature.
- Zuo, Y., & Li, W. (2024). *An Improved U-Net Lightweight Network for Semantic Segmentation of Weed Images in Corn Fields*. Jilin: Computers, Materials & Continua.