BENI SUEF UNIVERSITY
Faculty of Navigation Science & Space Technology

Graduation Project Report
2023
[Remote Sensing Solutions for Sustainable Resource Management Based on AI]

Prepared By
Mohammed Tawheed Abdulalwahab Amin

Sahar Abdulalim Himeda Mohammed

Neama Ahmed Ali Mohammed

Manar Marwan Atia Marwan

Mohammed Sobhy Mohammed Abdulalzeim

Khaled Elsayed Mohammed Amin

Belal Mousa Ouda Zourab

Fares Tarek Fawzy Mohammed

Mahmoud Hisham Shaaban Mohammed

Supervised By
Dr. MOHAMED MOURAD

24-7-2023

ABSTRACT

This graduation project thesis focuses on the fusion of remote sensing and artificial intelligence (AI) techniques for environmental analysis and infrastructure management. The thesis consists of interconnected projects that address data gathering and processing, water resources management, anomaly detection in crop patterns, crop classification, and oil storage detection.

The projects aim to automate the workflow of satellite imagery data processing, facilitating near-real-time access to ESA archive data. Utilizing remote sensing technology, comprehensive tools are developed for water resources management and drought monitoring. The thesis also explores anomaly detection in crop patterns and accurate crop classification, with a focus on rice crops in California. Additionally, the research investigates the detection and classification of oil storage facilities using high-resolution satellite data and AI algorithms.

The fusion of remote sensing and AI techniques throughout the projects enables efficient data analysis, informed decision-making, and optimized resource management. The outcomes provide practical tools and methodologies for researchers, policymakers, and industry professionals in diverse fields. The research contributes to sustainability, precision agriculture, and infrastructure management, leveraging the potential of advanced technologies.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to Professor Osama Shalabia "The Dean of the faculty" and for supporting us through our academic learning and our thesis advisor, [DR. Mohammed Mourad], for his guidance, support, and encouragement throughout the research project. His expertise, insights, and feedback have been invaluable in shaping the direction of the research and improving the quality of the thesis. We would also like to acknowledge the support provided by the funding organization [Nabta Playa], for their financial support throughout the research project. Their funding has enabled us to conduct the research and produce this thesis. We would like to thank the faculty members of the [Space Navigation Program] for their valuable contributions to our education and research. Their dedication to teaching and research has been a source of inspiration and motivation for us. We would like to acknowledge the support and encouragement from our families and friends. Their emotional support, encouragement, and understanding have been invaluable in helping us overcome the challenges and obstacles during the research. Finally, we would like to express our gratitude to all the participants who contributed to this research project. Their willingness to participate and provide valuable data and information has made this research possible. The completion of this thesis would not have been possible without the support and contributions of all these individuals and the funding organization. We are deeply grateful for their assistance and support.

DECLARATION

We hereby certify that this material, which We now submit for assessment on the program of study leading to the award of Bachelor of Science in (Space Navigation) is entirely our own work, that we have exercised reasonable care to ensure that the work is original, and does not to the best of our knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of our work.

Name	Signature
Mohamed Tawheed Abdulwaheb Amin	Mohammed Tawheed
Sahar Abdulalim Himeda Mohammed	Sahar Abdulalim
Neama Ahmed Ali Mohammed	Neama Ahmed
Manar Marwan Atia Marwan	Manar Marwan
Mohammed Sobhy Mohammed Abdulazeim	Mohammed Sobhy
Belal Mousa Ouda Zourab	Belal Mousa
Khaled Elsayed Mohammed Amin	Khaled Amin
Fares Tarek Fawzy Mohammed	Fares Tarek
Mahmoud Hisham Shaaban Mohammed	Mahmoud Hisham

TABLE OF CONTENT

1 List of Figures

List of figures CH.4

List of figures CH.5

List of figures CH.6

List of Figures CH 7

List of Figures CH 8

List of Figures GUI

II.LIST OF TABLES

III. LIST OF ACRONYMS/ABBREVIATIONS

GIS Geographic Information System

AI Artificial Intelligent

ESA	European Space Agency
APIs	Application Programming Interfaces
EO	Earth Observation
UTM	Universal Transverse Mercator
WGS	World Geodetic System
SAR	Synthetic-Aperture Radar
GRD	Ground Range Detected
MSI	Multispectral Instrument
VNIR	Visible and Near-Infrared
SWIR	Short Wave Infrared
WMS	Web Map Service
OGC	Open Geospatial Consortium
NDVI	Normalized Difference Vegetation Index
URL	Uniform Resource Locators
Tiff	Tag Image File Format
Json	JavaScript Object Notation
WFS	Web Feature Service
AWS	Amazon web Service
S3	Simple Storage Service
EPSG	European Petroleum Survey Group
GDAL	Geospatial Data Abstraction Library

AMSR-E Advanced Microwave Scanning Radiometer for EOS

JAXA Japan Aerospace Exploration Agency

JPL Jet Propulsion Laboratory

SMAP Soil Moisture Active Passive

SM Soil Moisture

SMAPEx SMAP Experiments

SMAPVEX Soil Moisture Active Passive Validation Experiment

L4_SM SMAP Level 4 Surface and Root Zone Soil Moisture (product/algorithm)

NASA National Aeronautics and Space Administration

NCCS NASA Center for Climate Simulation

NDMC National Drought Mitigation Center

NDVI Normalized Difference Vegetation Index

NESDIS National Environmental Satellite, Data, and Information Service

NOAA National Oceanic and Atmospheric Administration

YOLO You Only Look Once

CNN Convolutional Neural Network

AWS Amazon Web Services

ML Machine Learning

DL Deep Learning

SAM Segment Anything Model

1. INTRODUCTION

Remote sensing and artificial intelligence (AI) techniques have emerged as a promising approach for addressing complex environmental challenges and enhancing infrastructure management. This thesis focuses on the fusion of these two technologies to develop comprehensive tools for environmental analysis and infrastructure management. The thesis consists of five interconnected projects that address data gathering and processing, water resources management, anomaly detection in crop patterns, crop classification, and oil storage detection. The first project aims to automate the workflow of satellite imagery data processing, making ESA archive data accessible in near-real time. This project will develop comprehensive tools for data gathering and processing, streamlining the workflow of satellite data processing and analysis. The project will be implemented using the Python programming language and open-source libraries. The second project will develop comprehensive tools for water resources management and drought monitoring using remote sensing technology. This project will include the development of methodologies for drought monitoring, using AI algorithms to analyze satellite imagery data and identify areas of water stress. The third project will focus on crop anomaly detection in other crop types. This project will develop methodologies for identifying crop anomalies using remote sensing technology and deep learning algorithms. This project will seek to improve crop management practices and enhance yield prediction for these important crops. The project will be implemented using unmanned aerial systems and the TensorFlow library. The fourth project will investigate accurate crop classification, and yield prediction, with a focus on rice crops in California. This project will develop comprehensive tools for crop analysis, leveraging remote sensing technology and AI algorithms. This project will include the development of methodologies for crop classification, enabling more accurate and efficient crop management. The project will be implemented using the TensorFlow library for deep learning. The fifth project will examine the detection and classification of oil storage facilities using high-resolution satellite data and AI algorithms. This project will develop comprehensive tools for infrastructure management, enabling better decision-making and optimized resource management. The project will be implemented using Google Earth satellite imagery datasets. The fusion of remote sensing and AI techniques in these projects enables efficient data analysis, informed decision-making, and optimized resource management. The outcomes of these projects provide practical tools and methodologies for researchers, policymakers, and industry professionals in diverse fields. This research contributes to sustainability, precision agriculture, and infrastructure management, leveraging the potential of advanced technologies. In this thesis, we will explore the fusion of remote sensing and AI techniques for environmental analysis and infrastructure management. We will present five interconnected projects that address data gathering and processing, water resources management, anomaly detection in crop patterns, crop classification, and oil storage detection. The outcomes of these projects will provide practical tools and methodologies for researchers, policymakers, and industry professionals in diverse fields. This research will contribute to sustainability, precision agriculture, and infrastructure management, leveraging the potential of advanced technologies.

2.Project Description

The goal of this project is to develop a comprehensive framework for remote sensing applications, leveraging hyperspectral images obtained from the Sentinel satellite. Remote sensing plays a crucial role in gathering valuable information about specific areas and applications. Our framework will serve as a robust backbone, facilitating essential tasks such as data acquisition, region and time specification, and accounting for cloud coverage. By utilizing advanced algorithms and cutting-edge technology, we aim to enhance the effectiveness and efficiency of remote sensing applications.

The framework will begin with the development of algorithms specifically designed to download data from the Sentinel-1 and Sentinel-2 satellites. These algorithms will empower users to specify their desired regions of interest and timeframes for data collection, while considering the cloud coverage percentage. By leveraging the wealth of information provided by hyperspectral images, we can extract valuable insights to address various remote sensing needs.

Once the data is collected, a thorough preprocessing stage will be implemented. This stage will involve shaping and resizing the images to ensure they meet the specific requirements of the intended applications. By optimizing the data, we aim to improve the accuracy and reliability of subsequent analyses and modeling processes.

Due to the large volume of satellite images involved, advanced hardware infrastructure is essential. The size of individual tiles can range from 0.5 to 1 gigabyte, and covering the entire area of Egypt typically requires processing 25 to 30 tiles. To efficiently manage these computational demands, we will leverage cloud servers. Among the available options, we have chosen Amazon Web Services (AWS) as our preferred cloud solution. AWS offers scalability, reliability, and extensive features, making it an ideal choice for processing and storing vast amounts of remote sensing data.

The project workflow follows a systematic approach. It begins with data collection from diverse sources based on user-defined requirements. The collected data is then processed using our framework, ensuring it is appropriately shaped and sized for the specific applications. The framework will incorporate automation techniques to streamline the analysis and modeling processes.

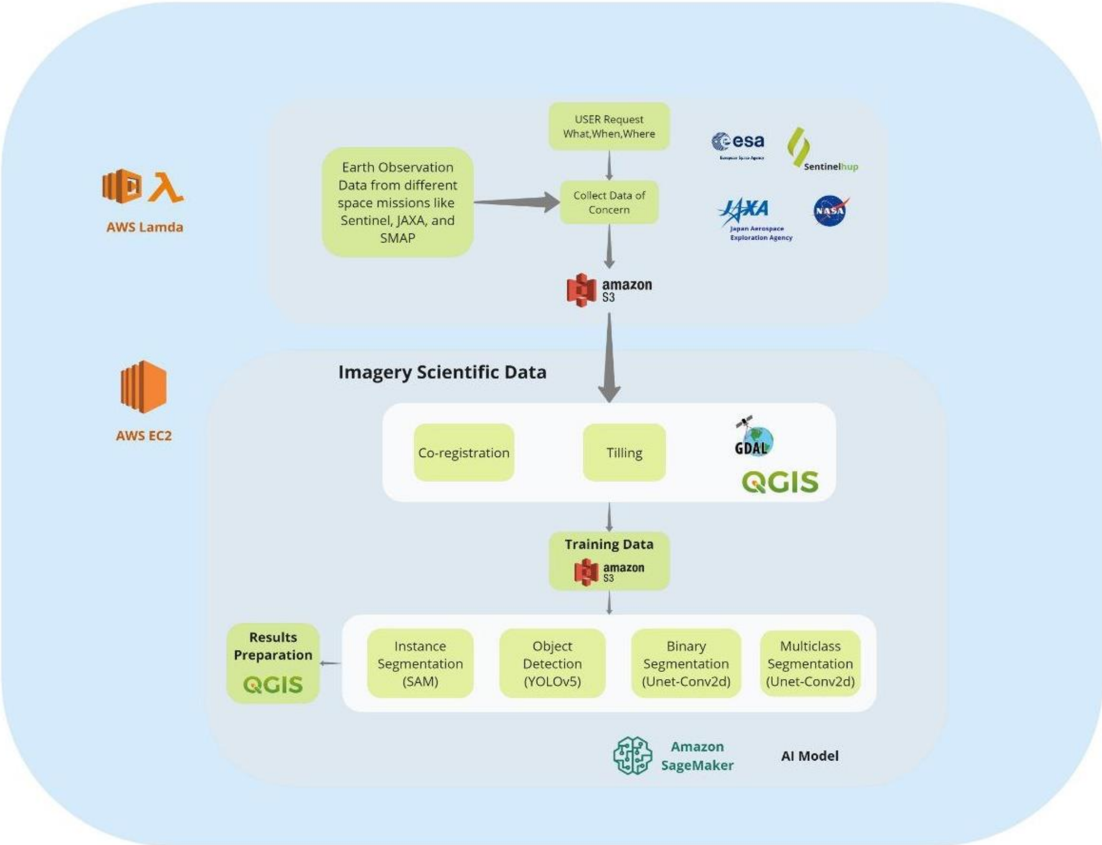


Figure 1-2

3.Project Impact

The projects described in this thesis have a significant impact on various aspects of sustainable development, aligning with Egypt's Vision 2030. Egypt's Vision 2030 is a comprehensive plan that aims to transform the country into a sustainable, knowledge-based economy, with a focus on inclusive growth, social justice, and environmental sustainability.

3.1. Advancing Egypt's Vision 2030

The projects described in this thesis have a significant impact on Egypt's Vision 2030, a comprehensive plan aimed at transforming the country into a sustainable and knowledge-based economy. Here's a summary of the specific impact on Egypt's Vision 2030

Direct Positive

1. Satellite imaging technology plays a crucial role in achieving these goals by enabling accurate monitoring and assessment of agricultural resources, helping to increase productivity and reduce food insecurity. It also supports urban planning and infrastructure development by providing detailed mapping and monitoring capabilities. Additionally, satellite images aid in disaster management and response, allowing for timely and effective relief efforts. Moreover, the technology facilitates improved access to education and healthcare services in remote areas through telemedicine and distance learning initiatives. Overall, satellite imaging technology enhances the quality of life in Egypt by addressing key socio-economic challenges and fostering sustainable development.

Indirect Positive

2. Satellite imaging technology contributes to justice and inclusion by promoting transparency and accountability in resource allocation and service delivery. It helps identify areas of inequality and enables targeted interventions to address disparities in access to resources and opportunities. By providing accurate data on population distribution and infrastructure gaps, satellite imaging supports evidence-based policymaking that ensures inclusive development across rural and urban areas. It also facilitates the monitoring and protection of marginalized communities, empowering them to participate in decision-making processes. Ultimately, satellite imaging technology strengthens justice, social inclusion, and participation by promoting equal rights, opportunities, and the sense of belonging to the Egyptian identity.

Direct Positive

3. Satellite imaging technology plays a vital role in fostering a strong economy by supporting informed decision-making in various sectors. It aids in assessing natural resources, such as mineral deposits and water sources, facilitating their sustainable utilization and contributing to economic diversification. Additionally, satellite images assist in monitoring infrastructure development and identifying potential investment opportunities. By providing accurate data on land use and urban planning, satellite imaging enables effective resource allocation and promotes a favorable business environment. Moreover, it helps in assessing environmental impacts and incorporating sustainable practices into economic development strategies. Overall, satellite imaging technology contributes to a competitive and diversified economy by enhancing knowledge-based growth, digital transformation, and sustainable resource management.

Indirect Positive

4. Satellite imaging technology supports knowledge and innovation by providing valuable data for scientific research and development initiatives. It enables researchers to study various aspects of the Earth's surface, climate patterns, and natural phenomena, fostering a better understanding of the world and driving innovation in various fields. Satellite images also aid in mapping and monitoring archaeological sites, contributing to the preservation of cultural heritage, and promoting historical research. By linking scientific research to education and development, satellite imaging technology facilitates the transfer of knowledge and encourages continuous learning. Ultimately, it enhances Egypt's capacity for knowledge-based development and fuels innovation across multiple sectors.

Direct Positive

5. Satellite imaging technology plays a critical role in promoting environmental sustainability by providing essential data for monitoring and addressing climate change impacts. It aids in assessing and managing natural resources, such as forests, water bodies, and biodiversity, contributing to their conservation and sustainable use. Satellite images help in identifying areas prone to natural disasters, facilitating effective disaster preparedness and response strategies. Additionally, satellite technology supports the expansion of renewable energy sources by assisting in site selection and monitoring energy infrastructure. By promoting sustainable consumption and production patterns, satellite imaging technology contributes to creating an integrated and sustainable ecosystem for the present and future generations.

Indirect Positive

6. Satellite imaging technology strengthens governance by providing reliable and objective data for evidence-based decision-making. It aids in monitoring and evaluating the performance of state institutions, private sector entities, and civil society organizations. By promoting transparency and accountability, satellite images contribute to combating corruption and ensuring adherence to laws and regulations. The technology also supports the enforcement of the rule of law by providing accurate information for legal proceedings and investigations. Ultimately, satellite imaging enhances the efficiency and effectiveness of governance in Egypt by establishing an institutional framework that fosters transparency, accountability, and the rule of law.

Direct Positive

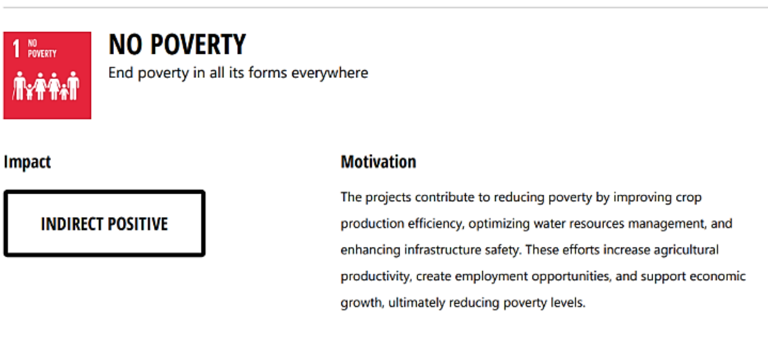
7. Satellite imaging technology plays a crucial role in enhancing Egyptian peace and security by supporting various aspects of national and regional security. It aids in monitoring and securing borders, providing valuable data on border activities and potential security threats. Satellite images also contribute to counter-terrorism efforts by assisting in the identification and tracking of terrorist activities and networks. Additionally, satellite technology facilitates the monitoring and protection of critical infrastructure, ensuring sustainable energy security and information security (cyber). By providing accurate data on natural resources, satellite imaging helps in ensuring food and water security and supporting environmental stability. Overall, satellite imaging technology strengthens Egyptian peace and security by addressing key security challenges and fostering stability and sustainable development.

Indirect Positive

8. Satellite imaging technology strengthens Egypt's leadership position by enhancing its capabilities in regional and international partnerships. By providing accurate and up-to-date data, satellite images contribute to evidence-based decision-making and facilitate collaboration on regional development initiatives. The technology enables Egypt to showcase its expertise and resources in areas such as infrastructure development, environmental conservation, and disaster management, solidifying its leadership role in these domains. Satellite imaging also supports information sharing and cooperation with other countries, fostering trust and building stronger partnerships. Ultimately, satellite imaging technology reinforces Egypt's leadership position by enabling effective collaboration and showcasing its commitment to comprehensive development at the regional and international levels.

3.2. SDGs by Impact assessment Tool

This research assesses the impact of integrated remote sensing and AI projects on the Sustainable Development Goals (SDGs). By utilizing advanced technologies, these projects contribute to sustainable development in various areas. The assessment explores how these projects align with specific SDGs, providing valuable insights into their transformative potential for achievable sustainability targets.



ZERO HUNGER

End hunger, achieve food security and improved nutrition and promote sustainable agriculture

Impact

DIRECT POSITIVE

Motivation

The projects contribute to reducing poverty by improving crop production efficiency, optimizing water resources management, and enhancing infrastructure safety. These efforts increase agricultural productivity, create employment opportunities, and support economic growth, ultimately reducing poverty levels.

GOOD HEALTH AND WELL-BEING

Ensure healthy lives and promote well-being for all at all ages

Impact

INDIRECT POSITIVE

Motivation

While not the primary focus, the projects indirectly promote good health and well-being. By ensuring safe water resources, enhancing infrastructure safety, and optimizing agricultural practices, they create healthier environments, reducing health risks and supporting overall well-being.

QUALITY EDUCATION

Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all

Impact

INDIRECT POSITIVE

Motivation

The projects provide opportunities for knowledge sharing, capacity building, and research in the fields of remote sensing, AI, and agriculture. By promoting advanced technologies and methodologies, they contribute to quality education and skill development for researchers, professionals, and policymakers.

GENDER EQUALITY

Achieve gender equality and empower all women and girls

Impact

INDIRECT POSITIVE

Motivation

The projects indirectly contribute to gender equality by promoting equal access to resources and opportunities in agriculture and infrastructure management. By empowering women in these sectors, they aim to reduce gender disparities and promote gender equality.

CLEAN WATER AND SANITATION

Ensure availability and sustainable management of water and sanitation for all

Impact

DIRECT POSITIVE

Motivation

The projects align with SDG 6 by monitoring and managing water resources. They optimize water usage in agriculture, ensure the safety of water storage facilities, and promote responsible water resource management, contributing to clean water and sanitation goals.

AFFORDABLE AND CLEAN ENERGY

Ensure access to affordable, reliable, sustainable and modern energy for all

Impact

INDIRECT POSITIVE

Motivation

While the direct impact may be limited, the projects indirectly support SDG 7 by optimizing energy usage in agricultural practices and enhancing infrastructure efficiency. These efforts contribute to affordable and clean energy consumption.

DECENT WORK AND ECONOMIC GROWTH

SUSTAINABLE CITIES AND COMMUNITIES

Make cities and human settlements inclusive, safe, resilient and sustainable

Impact

INDIRECT POSITIVE

Motivation

While not directly related to urban areas, the projects contribute to sustainable development by enhancing infrastructure safety, promoting efficient resource management, and supporting the development of resilient communities. They aim to create safe, inclusive, and sustainable cities and communities.

RESPONSIBLE PRODUCTION AND CONSUMPTION

Ensure sustainable consumption and production patterns

Impact

INDIRECT POSITIVE

Motivation

The projects align with SDG 12 by promoting responsible consumption and production. They optimize resource management, reduce waste, and promote sustainable agricultural practices, contributing to responsible consumption and production patterns.

CLIMATE ACTION

Take urgent action to combat climate change and its impacts

Impact

DIRECT POSITIVE

Motivation

The projects directly support SDG 13 by monitoring and managing water resources, optimizing energy usage, and enhancing infrastructure safety. By contributing to climate change mitigation and adaptation efforts, they align with climate action goals.

PARTNERSHIPS FOR THE GOALS

Strengthen the means of implementation and revitalize the global partnership for sustainable development

Impact

DIRECT POSITIVE

Motivation

The projects emphasize collaboration between different stakeholders, including researchers, policymakers, and industry professionals. By fostering effective partnerships, knowledge sharing, and capacity building, they support the implementation of the SDGs and contribute to achieving sustainable development goals.

LIFE ON LAND

Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss

Impact

INDIRECT POSITIVE

Motivation

Through optimized resource management, responsible agricultural practices, and infrastructure safety, the projects aim to preserve terrestrial ecosystems, protect biodiversity, and promote sustainable land use. They contribute to achieving the goals of SDG 15.

PEACE, JUSTICE AND STRONG INSTITUTIONS

Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels

Impact

INDIRECT POSITIVE

Motivation

The projects indirectly contribute to SDG 16 by fostering transparency, accountability, and safety in infrastructure management. By promoting good governance, they aim to support peace, justice, and the development of strong institutions.

PARTNERSHIPS FOR THE GOALS

Strengthen the means of implementation and revitalize the global partnership for sustainable development

Impact

DIRECT POSITIVE

Motivation

The projects emphasize collaboration between different stakeholders, including researchers, policymakers, and industry professionals. By fostering effective partnerships, knowledge sharing, and capacity building, they support the implementation of the SDGs and contribute to achieving sustainable development goals.

In summary, the integrated remote sensing and AI projects outlined in this research have a profound impact on both the Sustainable Development Goals (SDGs) and Egypt's Vision 2030. These projects contribute to various SDGs, including economic growth, food security, sustainable agriculture, water resource management, infrastructure development, innovation and technology, environmental sustainability, and collaborative partnerships. By promoting economic prosperity, ensuring food self-sufficiency, advancing sustainable agricultural practices, optimizing resource management, developing resilient infrastructure, fostering innovation, and supporting environmental conservation, these projects align with Egypt's Vision 2030 goals. Through their collective efforts, they contribute to the realization of sustainable development objectives, paving the way for a prosperous and sustainable future in Egypt.

4.DATA GATHERING

Abstract

The fields of Remote Sensing and GIS have become increasingly important and attractive due to expanding opportunities and growing interest in better understanding our environment. In the context of prototyping a remote sensing solution that can collect and analyze data about the Earth's surface from a distance using sensors and imaging technology, it can provide valuable insights into various aspects of the Earth's surface, such as land cover, vegetation, water resources, and urban development, using machine learning and AI algorithms. The process of gathering and processing satellite imagery datasets can be time-consuming and resource-intensive, requiring significant manual effort to download, organize, and prepare the data for analysis. Therefore, the main focus of this thesis is to investigate the potential for the process of satellite imagery dataset gathering and post-processing to develop an automated workflow that can download, process, and prepare satellite data for AI models in near-real-time. To achieve this goal, the thesis will provide an automated workflow that is designed to perform mainly four sequential tasks. The first task is data downloading, where the workflow can automatically download satellite imagery data from the ESA archive using Sentinel Hub APIs based on user input such as the area of interest and date range. If there is a well-organized data system, access to satellite data will be more efficient and productive, facilitating the analysis process by providing a unified view of the data. The second task is to automatically organize the downloaded data in a structured manner to provide a flexible and smooth approach to obtaining high-quality data. If the images used in the analysis are acquired from different sensors, it is difficult for any AI model to deal with them because each sensor has unique characteristics, such as different spatial resolutions, viewing angles, and spectral ranges. Therefore, the third step is to tackle this problem, where the co-registration process will be done automatically to align the images spatially and have the same scale, orientation, and geometrical features, which are essential for accurate analysis and interpretation of the data. Afterwards, tiling is always needed as large images have to be split into tiles before being used by the AI model and this is what the final step does. After these steps, the data will be ready to be passed into the AI models. Designing this automated workflow required careful planning and coordination between different components of the system to be flexible and adaptable to use cases so that it can handle different types of data, reduce the time and effort required to process these data, and enable faster and more accurate analysis of the data.

Introduction

1.1 Background

Artificial satellites have been in existence for over 50 years since the launch of Sputnik in 1957. This launch ushered in the space age and confirmed Sir Isaac Newton's theoretical explanation of how an artificial satellite could be launched into Earth orbit. Today, the world of satellites can be categorized into two broad areas: scientific satellites and applications satellites. Scientific satellites explore our world, our solar system, our galaxy, and radiation from the Van Allen Belts to cosmic radiation. On the other hand, applications satellites provide practical services to people here on Earth.

These satellites are a part of our everyday lives whether we know it or not. One of these applications satellites is remote sensing and Earth observation satellites that truly serve humankind. [**1**]

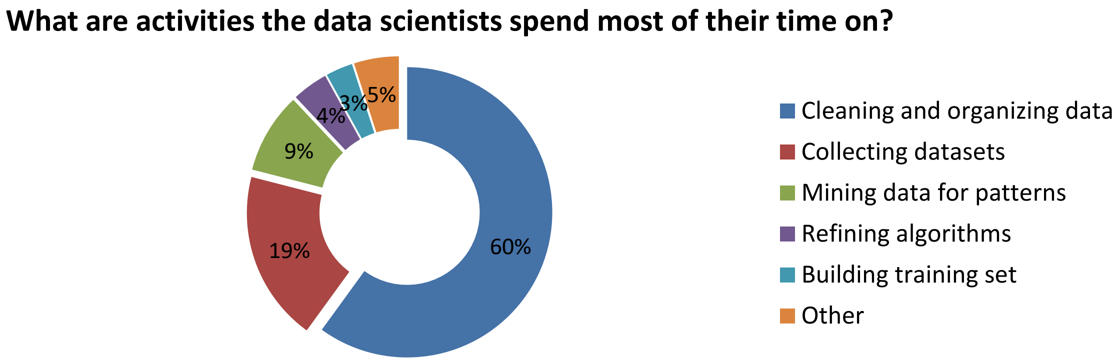
Remote sensing via satellite has become a key service that is used in many civil applications such as agriculture, forestry, mining (and prospecting for many types of resources), map making, research in geosciences, urban planning, land speculation, disaster warning, and other activities to sustain the biodiversity of plant and animal life on our planet. **[1]** Responding to major disasters routinely involves analysis of satellite imagery, which means that they have become more sophisticated and reliable. So, providing an automation process of satellite imagery analysis allows for more monitoring and fast response to some issues. **[2]**

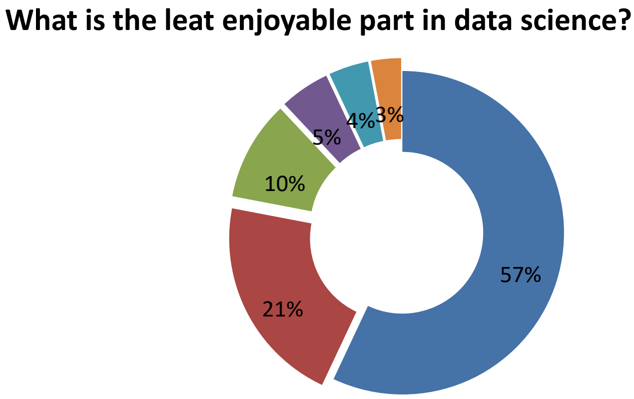
The process of data collection is a major bottleneck in machine learning and an active area of research in multiple communities. The emergence of new machine learning applications and the use of deep learning techniques that require larger amounts of labeled data have made data collection a critical issue. Researchers in machine learning, natural language, computer vision, and data management communities are actively working on data collection. The integration of machine learning and data management for data collection is part of the trend of big data and Artificial Intelligence integration, which presents many opportunities for new research.**[2]**

Remote sensing data processing has significant societal value in applications such as urban monitoring, fire detection, and flood prediction. The process of remotely sensed multispectral or radar images is crucial for addressing economic and environmental issues. Remote sensing has become a multidisciplinary field that relies heavily on machine learning and signal processing algorithms to efficiently process acquired data and provide accurate results.**[3]**

1.2 Problem statements

In the charts above, a survey of data scientists found that they spend the majority of





CH 4 Figure 1: Top: Time consuming activities of data science, Bottom: Least enjoyable activities of data science **[4]**

CH 4 Figure 2: Example of a multispectral image acquired by the passive Sentinel2-L1C sensor over east of Beni Suef, Egypt. The image is obtained from sentinelhup request builder platform. The image is taken on 27-6-2023: spectral channels corresponding to blue (a), green (b), red (c), and NIR radiation (d); the true color composite (e), in which the R, G, and B components of the displayed image are associated with the red, green, and blue channels of the multispectral image, respectively; and a false color composite (f), in which the R, G, and B components of the displayed image are associated with the NIR, red, and green channels of the multispectral image, respectively.CH 4 Figure 3: Top: Time consuming activities of data science, Bottom: Least enjoyable activities of data science **[4]**

their time cleaning and organizing data 60%, followed by collecting datasets 19%. This means that around 80% of their time is spent on preparing and managing data for analysis. Data preparation is viewed as the least enjoyable part of their work by 76% of data scientists, while 57% find cleaning and organizing data to be the least enjoyable, and 19% say the same about collecting datasets.**[4]** Therefore, this thesis comes to solve two major challenges.

First challenge: Due to many satellites in orbit, According to the UCS there were 971 EO satellites in orbits on the 30th April 2021**[5],** massive data are added day by day to archives. So, it consumes time for non-expert user to deal with such big data to download the correct images for best use.

Second challenge: the satellite imagery acquired at different times or by different sensors can result in misalignments or distortions in the images which are not match each other in terms of their spatial orientation and scale. A certain process must be done to align these images to be acceptable by models for train or test.

1.3 Aims of the study.

Concerning the problem statements discussed above, the overall aim of this thesis is to build an automatic process beginning downloading the images according to user request from the area of interest and time interval of images, to solve the first challenge discussed above, passing to co-register the downloaded images to align each other making the images more suitable to use, to solve the second challenge discussed above, finally split these images into patches based on the user request for the height and weight for each patch. Providing by this workflow, data is ready to be processed by some AI technique directly.

2 Literature review

2.1 Remote sensing history

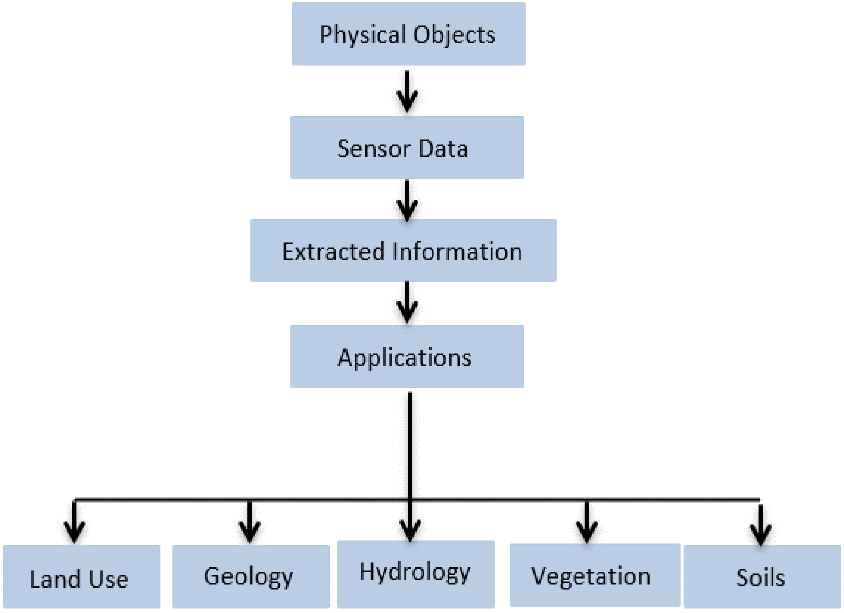
Remote sensing has a rich history, dating back to the use of cameras and the naked eye from balloons, kites, and pigeons. Advancements in aircraft and technology during wars led to aerial remote sensing capabilities, while the space age brought about space remote sensing, which has been used for military reconnaissance, weather prediction, and environmental analysis. Today, many countries operate their own remote sensing satellite systems for various applications. The first civil remote sensing satellite was launched in 1960, followed by the Landsat program in 1972. France's SPOT program was successful in producing remote sensing data for commercial use. Currently, over 30 nations operate various types of satellite systems, with the latest being hyper spectral imaging sensors, which provide narrow-band imaging and present challenges for data processing. Remote sensing technology has also played a significant role in post-Cold War era stability and served as the basis for robotic exploration of other planets.**[1]**

2.2 Remote sensing definition

The field of remote sensing has been defined in various ways. One of the most specified definitions is “The science of deriving information about an object from measurements made at a distance from the object without actually coming in contact with it.” **[10]**. The act of remote sensing involves detecting and recording energy that is reflected or emitted, and then analyzing and utilizing that information **[11].**

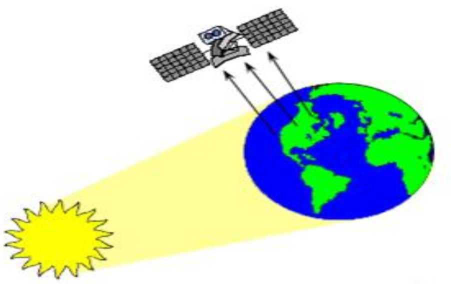
2.3 Remote sensing process

To achieve the process of remote sensing, radiation is sent from a satellite or a natural radiation source like the sun towards the Earth or a specific object of interest. As this radiation reaches the object, it reflects back energy. This reflected energy is captured by remote sensing satellites and transmitted to a remote station for conversion into images. Remote sensing involves gathering data in the form of images that reflect energy, providing more spectral and spatial information compared to regular images. This information reveals the structure, shape, and texture of the object being studied which are processed to extract information about the object, such as water bodies, agriculture, land cover classification, and more.**[12]**

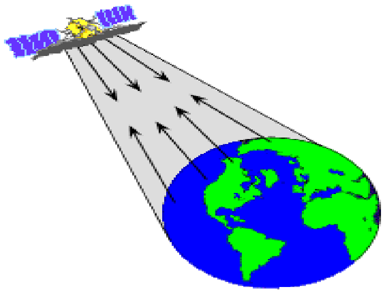


CH 4 Figure 2: Overview of the remote sensing process [10].

There are two types of remote sensing: passive sensing, which measures energy generated by natural sources such as the sun using passive sensors, and active sensing, which measures energy generated by satellite sources such as electromagnetic radiation, radar, and microwave which are called active sensors**[13].**



CH 4 Figure 3: Passive Sensing

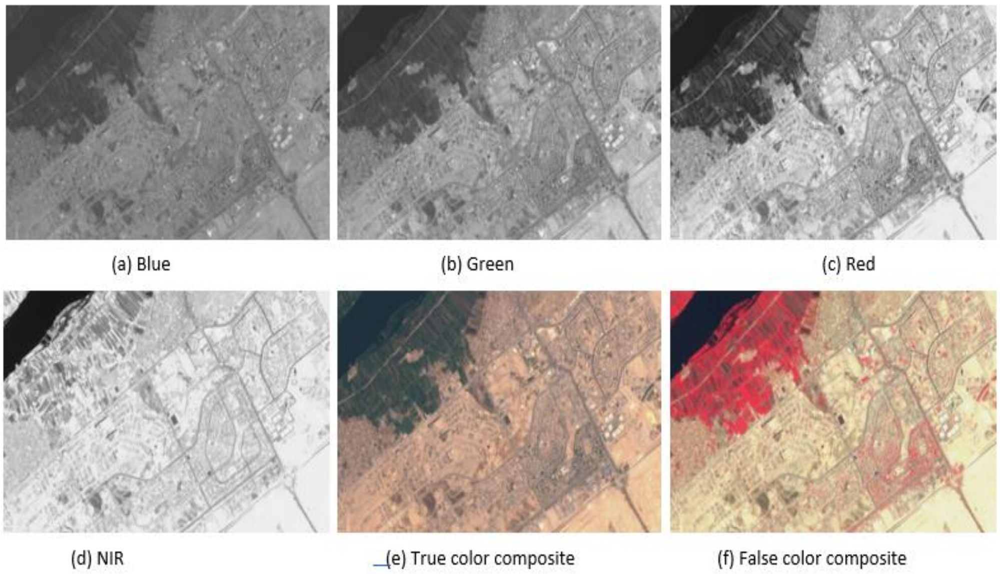


CH 4 Figure 4: Active Sensing

2.3.1 Passive Sensors for EO

Passive sensors receive electromagnetic radiation from a portion of the Earth's surface, either from the reflection of incident solar radiation or from spontaneous emission by the surface itself. The physical quantity measured by a passive EO sensor is the spectral radiance, which represents the power per unit wavelength that travels in a unitary solid angle centered on a given direction through a unitary surface. The received radiance depends on the reflective properties of the observed surface in visible (i.e., with a wavelength between approximately 0.4 and 0.7µm), near-infrared range (NIR, 0.7–1.1µm), and short-wave infrared range (SWIR, 1.1–1.35µm, 1.4–1.8µm, and 2–2.5µm).

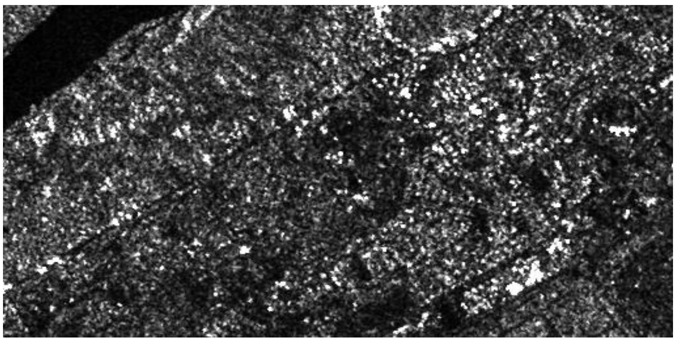
Passive EO sensors are usually multispectral, as shown in **Figure 5**, and hyperspectral sensors collect data from hundreds of channels with narrow bandwidths. Multispectral acquisition can be accomplished using prisms and optical filters, or separate cameras that operate in distinct wavelength ranges. Data collected by passive sensors are affected by atmospheric and Sun-illumination conditions**. [23]**



CH 4-Figure 5: Example of a multispectral image acquired by the passive Sentinel2-L1C sensor over east of Beni Suef, Egypt. The image is obtained from sentinelhup request builder platform. The image is taken on 27-6-2023: spectral channels corresponding to blue (a), green (b), red (c), and NIR radiation (d); the true color composite (e), in which the R, G, and B components of the displayed image are associated with the red, green, and blue channels of the multispectral image, respectively; and a false color composite (f), in which the R, G, and B components of the displayed image are associated with the NIR, red, and green channels of the multispectral image, respectively.

2.3.2 Active Sensors for EO

Active sensors for EO transmit an electromagnetic pulse towards the Earth's surface and receive the resulting "echo" signal. Microwave signals are typically used for 2D remote sensing image acquisition based on a radar (Radio Detection and Ranging) instrument. Synthetic Aperture Radar (SAR) uses the motion of the platform to simulate a long antenna and achieve high spatial resolution along the flight direction. Active sensors are almost insensitive to cloud cover and atmospheric conditions and provide day-and-night and all-weather acquisition capability. **[23]**



CH4 Figure6: Example of a SAR image acquired by the Sentinel1 over east of Beni Suef, Egypt. The image is obtained from sentinel hub request builder platform. The image is taken on 24-6-2023: channel corresponding to polarization.

2.4 Articles reviews

The power of satellite imagery shines in how far it can be used in the service of humans and how far the researchers and scientists enable from analyzing satellite imagery in their way of human evolution. The following are some articles in which research mainly relied on satellite imagery datasets to develop their research and came out with accurate results helping in building a more powerful community.

2.4.1 Review about article1

In an article about the challenges of accurately mapping paddy rice in cloudy and foggy regions using optical images. An automatic mapping method using Sentinel-1 Synthetic Aperture Radar (SAR) data pre-processes by The Sentinel Applications Platform (SNAP) 6.0.0 and Python 3.8.2 versions. The method, called Auto-CFM, is based on constrained feature matching, and can handle partial missing images. The study validated this method in different provinces with varying planting times, climates, and topographies. The article concludes that this method can provide timely and accurate information on rice cultivation areas before harvest, which is essential for coordinating agricultural production and ensuring grain security. **[6]**

2.4.2 Review about article2

In research about the generation of national agricultural land cover maps yearly taking into its account the varying environmental conditions. The study uses a random forest classifier and dense time series data from Sentinel-2 and Landsat 8 in combination with monthly Sentinel-1 composites and environmental data to evaluate the relative importance of optical, radar, and environmental data. The study concludes that the integrated use of optical time series and SAR data, along with variables describing local and seasonal environmental conditions, can effectively map large-area crop types on an annual basis. **[7]**

2.4.3 Review about article3

A literature article highlights the importance of monitoring shorelines using satellite images to identify environmental issues such as coastal erosion. The article proposes a technique for automatically creating labeled datasets of satellite images for shoreline detection by integrating data from satellite photos and publicly accessible shoreline data. The study uses Sentinel-2 data for creating a dataset for shoreline detection due to its better performance in terms of spatial resolution and revisits period compared to other public continuous Earth observation missions. **[8]**The practical implementation involves obtaining Sentinel-2 Level-1C tiles using the Plateforme d'Exploitation des Produits Sentinel of the Centre National d'études Spatiales **[9]**, selecting satellite images based on location and date, and processing the tiles to extract semantically annotated samples. This involves generating Level-2A products, projecting shoreline paths, and splitting the tiles into sub-tiles of size 256x256. The processing focuses on sub-tiles containing shoreline paths and involves creating a binary segmentation map for each sub-tile. The article concludes that the proposed method eliminates the need for manual annotation and subjective interpretation of satellite images and can be used to train neural models for sea-land segmentation. **[8]**

It is obvious that satellite imagery plays a crucial role in a wide range of applications and research areas beyond those mentioned earlier. Its importance as a primary source of information is evident in the numerous fields that rely on it for accurate and up-to-date data.

3. Methodology

3.1 Dataset

As shown previously, the literature review ensures that the Sentinel satellites are effective in many aspects and service many applications so this thesis limits satellite imagery datasets only to Sentinel1 and Sentinel2 images.

The algorithm is built mainly by Sentinel Hub which is a satellite imagery service that utilizes multi-spectral and multi-temporal data products and supports various data collections, which include freely available and commercial collections to provide access to large volumes of remote sensing information and related EO. The service is capable of fully automated archiving, real-time processing, and distribution of data. By using APIs, an intermediary that allows applications to access the features or data of another application or system, users can retrieve satellite data for their area of interest (AOI) and specific time range from the complete archives in just a matter of seconds. **[14]**

3.1.1 Background sentinel1 [15]

3.1.1.1 Overview

The Sentinel-1 mission is a constellation of two polar-orbiting satellites with a planned launch of Sentinel-1A on April 3, 2014, and Sentinel-1B on April 25, 2016. The satellites share the same orbit plane with a 180° orbital phasing difference.

CH 4 Figure 7: Sentinel1 SAR

3.1.1.2 Mission Orbit

- Sun-synchronous, near-polar, circular orbit
- 693 km orbit height
- 98.18° inclination
- 12-day repeat cycle at Equator with one satellite, 175 orbits/cycle.

3.1.1.3 Instrument Payload

C-band Synthetic Aperture Radar

- Centre frequency: 5.405 GHz
- Polarization: VV+VH, HH+HV, HH,VV

3.1.1.4 Mission Objectives

The mission objectives include land monitoring of forests, water, soil, and agriculture, emergency mapping support in the event of natural disasters, marine monitoring of the maritime environment, sea ice observations and iceberg monitoring, production of high-resolution ice charts, forecasting ice conditions at sea, mapping oil spills, sea vessel detection, and climate change monitoring.

3.1.1.5 Acquisition Modes
Sentinel-1 operates in four exclusive acquisition modes:

- Stripmap (SM) for small islands
- Interferometric Wide swath (IW) for land
- Extra-Wide swath (EW) for wide area coastal monitoring
- Wave (WV) for open ocean

3.1.1.6 Resolution
Spatial resolutions depend on the acquisition mode and the level of processing.

High resolution Level-1 GRD

Resolution -> 20x22 m

Pixel spacing -> 10x10 m.

3.1.2 Background sentinel2

3.1.2.1 Overview [17]
The Copernicus SENTINEL-2 mission comprises a constellation of two polar-orbiting with planned launch SENTINEL-2A on 23 June 2015 and Sentinel-2B on 7 March. The satellites share the same orbit plane with a 180° orbital phasing difference.

Figure 8: Sentinel2 MSI CH4

3.1.2.2 Mission Orbit [17]

- Sun-synchronous, near-polar, circular orbit 786 km orbit height
- 98.62 ° inclination
- 10-day repeat cycle, 143 orbits/cycle.

3.1.2.3 Instrument Payload [17]
Each of the satellites carries a single payload.

- The Multi-Spectral Instrument measures the Earth's reflected radiance in 13 spectral bands from VNIR to SWIR.

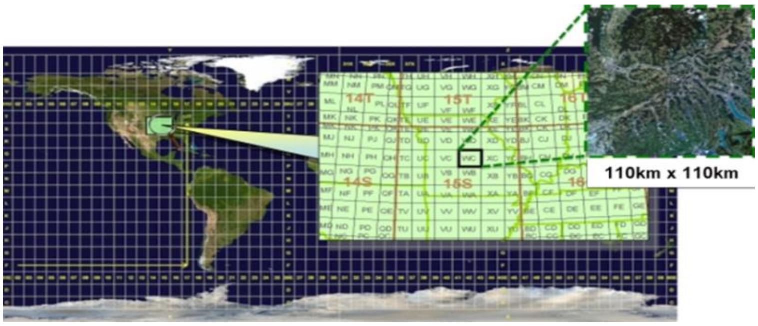
3.1.2.4 Mission Objectives
The mission provides systematic global high-resolution multispectral imagery with a high revisit frequency, enhanced continuity of multi-spectral imagery provided by the SPOT series, and observations for the next generation of operational products such as land-cover maps, land-change detection maps, and geophysical variables. **[16]**

3.1.2.5 Product types

Type	Code	Description	Users	Production & Distribution
User Product	Level-1B	Top-Of-Atmosphere radiances in sensor geometry	Expert Users	Systematic generation and online distribution
	Level-1C	Top-of-atmosphere reflectances in cartographic geometry	All Users	Systematic generation and online distribution
	Level-2A	Atmospherically corrected Surface Reflectances in cartographic geometry		

CH4 Table 1 : Sentinel2 Product Types [17]

For Level-1C and Level-2A, the granules, also called tiles, are 110x110 km2 ortho-images in **UTM/WGS84** projection.



CH 4 Figure 9: Level1-1C Product tiling [17]

3.1.2.6 Resolution

The spatial resolution of SENTINEL-2 is dependent on the particular spectral band.

Name	Description	Resolution
B01	Coastal aerosol, 442.7 nm (S2A), 442.3 nm (S2B)	60m
B02	Blue, 492.4 nm (S2A), 492.1 nm (S2B)	10m
B03	Green, 559.8 nm (S2A), 559.0 nm (S2B)	10m
B04	Red, 664.6 nm (S2A), 665.0 nm (S2B)	10m
B05	Vegetation red edge, 704.1 nm (S2A), 703.8 nm (S2B)	20m
B06	Vegetation red edge, 740.5 nm (S2A), 739.1 nm (S2B)	20m
B07	Vegetation red edge, 782.8 nm (S2A), 779.7 nm (S2B)	20m
B08	NIR, 832.8 nm (S2A), 833.0 nm (S2B)	10m
B8A	Narrow NIR, 864.7 nm (S2A), 864.0 nm (S2B)	20m
B09	Water vapour, 945.1 nm (S2A), 943.2 nm (S2B)	60m
B10	SWIR – Cirrus, 1373.5 nm (S2A), 1376.9 nm (S2B)	60m
B11	SWIR, 1613.7 nm (S2A), 1610.4 nm (S2B)	20m
B12	SWIR, 2202.4 nm (S2A), 2185.7 nm (S2B)	20m

CH4 Table 2 : Sentinel2 available bands [18]

3.2 Access Sentinel Hub images using WMS.

To avoid the complexities of Sentinel satellites data and without need for large storage volumes and lots of processing power. OGC using the standard WMS, One of API of Sentinel Hub, not only provides access to raw satellite data but also to processed products such as true color imagery and NDVI. Access to the service is done via a custom server instance URL. **[20]**

There are some parameters needed to involve in WMS request shown in the table below.

WMS parameter	Information
data collection Required	Type the mission which to return the results.
bbox Required	The bounding box of the requested image must be defined by four coordinates in the specified coordinate reference system. The coordinates should be separated by commas and represent the top-left and bottom-right corners of the bounding box.
time Optional	The time for which to return the results. It may be single time or time interval. The result is based on all scenes between the specified times. Note: Requesting a single value for TIME parameter is deprecated. Sentinel Hub interpreted it as a time interval [given time - 6 months, given time]. For the vast majority of cases this resulted in unnecessary long processing time thus we strongly encourage you to always use the smallest possible time range instead. The default: none (the last valid image is returned).
CRS Optional	The coordinate reference system (CRS) in which the bounding box (BBOX) is defined and in which the resulting image will be returned. The default: "EPSG:3857".
format Optional	The returned image formats. The default: "image/png", other options: "image/jpeg", "image/tiff".
width Required	The returned image width in pixels. Optional when HEIGHT is used
height Optional	The returned image height in pixels. Required, unless WIDTH is used.
layer Required	The preconfigured layer (image) to be returned as true color, NDVI and etc.
maxcc Optional	The maximum cloud coverage of returned images.

CH4 Table **3:** Standard common WMS parameters [20]

So, the question arises here, how does user configure a layer?

1. Sign up for the sentinel Hub account.
2. Create a new configuration for layers.
3. Add a new layer then choose from available options which have to fit with the use case.
4. With configuration ID and layer ID, the sentinel images can be accessed from any Sentinel Hub APIs.

The workflow uses WMS to obtain Sentinel1 and Sentinel2 images with different characteristics of images as will be shown later.

3.2.1 Characteristics of the returned images with WMS

The requests of WMS return two outputs to each image.

1. Image based on the request with
   - extension -> tiff
   - size -> 512 width
2. Json file with valuable information about the image.

3.3 Access Sentinel Hub images using WFS.

Another standard of OGC is WFS. It provides access to the geometric (vector) metadata about the available data collection tiles **[21].** The advantage of this service is that Sentinel-2 imagery and additional data are stored on AWS S3 storage buckets, which are more stable of this workflow as it uses AWS to run the request and to store the downloaded data.

The data at AWS is the same as original S-2 data provided by ESA. The archive of Sentinel-2 data at AWS consists of two buckets, one containing L1C and the other containing L2A data **[22].** But the available archive up till now is that one for sentinel2-L1C data, so the algorithm is built only with this type of data in this service.

There are some parameters needed to be involved in WFS request shown in the table below.

WSF parameter	Information
data collection Required	Type the mission which to return the results. Here it must be -> DataCollection.SENTINEL2_L1C.
bbox Required	The bounding box of the requested image must be defined by four coordinates in the specified coordinate reference system. The coordinates should be separated by commas and represent the top-left and bottom-right corners of the bounding box.
time Optional	The time for which to return the results. The default: none (the last valid image is returned).
CRS Optional	The coordinate reference system (CRS) in which the bounding box (BBOX) is defined and in which the resulting image will be returned. The default: "EPSG:3857".
maxcc Optional	The maximum cloud coverage of returned images.

CH4 Table 4 : Table Standard WFS parameters [21]

Another request must follow WFS request. It is AwsTileRequest

AwsTile parameter	Information
data collection Required	Type the mission which returns the results. Here it must be -> DataCollection.SENTINEL2_L1C.
time Optional	Sentinel-2 tile time can be accessed from ESA tile ID which get from the result of WFS.
tile_name Required	Sentinel-2 tile_name can be accessed from ESA tile ID which get from the result of WFS.
aws_index Required	the last number in tile AWS path which can be accessed from ESA tile ID which get from the result of WFS.
Bands Optional	Specifying bands among 13 available bands of sentinel2-L1C. The default: all 13 bands -> [“B01”, “B02”, “B03”, “B04”, “B05”, “B06”, “B07”, “B08”, “B8A”, “B09”, “B10”, “B11”, “B12”].
metafiles Optional	Specifying among available metafiles of sentinel2-L1C The default: none.
maxcc Optional	The maximum cloud coverage of returned images.

CH4 Table 5: Aws Tile parameters [22]

3.3.1 Characteristics of the returned tiles with WFS

The requests of WFS return to each tile.

1. The selected bands based on the request with
    - extension -> jp2
    - size -> 5490X5490
2. The selected metafiles

3.3.2 Disadvantage of request with WFS

According to the input coordinates of the area of interest, the request tries to get the tile for this given coordinate and if there is not, it tries to get the tile of the closest coordinate that is unlike the request with WMS, which returns the image of the selected coordinate. To solve this problem, it’s recommended to co-register of the obtained images of WFS request with images of WMS request. So, if the tiles as required, the images exist and if the tiles are different than the required, it returns black images, without information, as they are spatially different.

3.4 The algorithm

The algorithm consists of two main parts.

1. Data Collection
    a. Request one type of data collection
    b. Request two types of data collection

2. Data Preparation
    a. Data Co-registration
    b. Data tilling

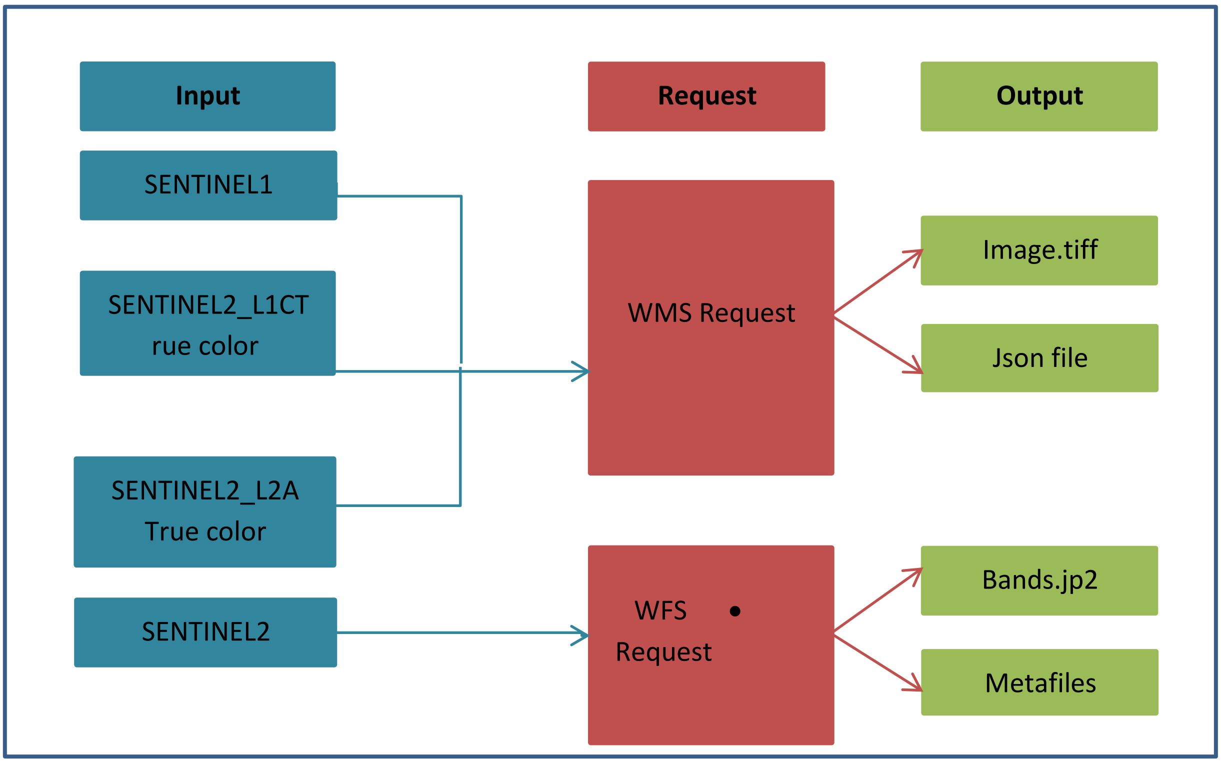
3.4.1 Data Collection

The algorithm provides two flows to cover most of the use cases.

3.4.1.1 Request one type of data collection

This algorithm includes the following collection:

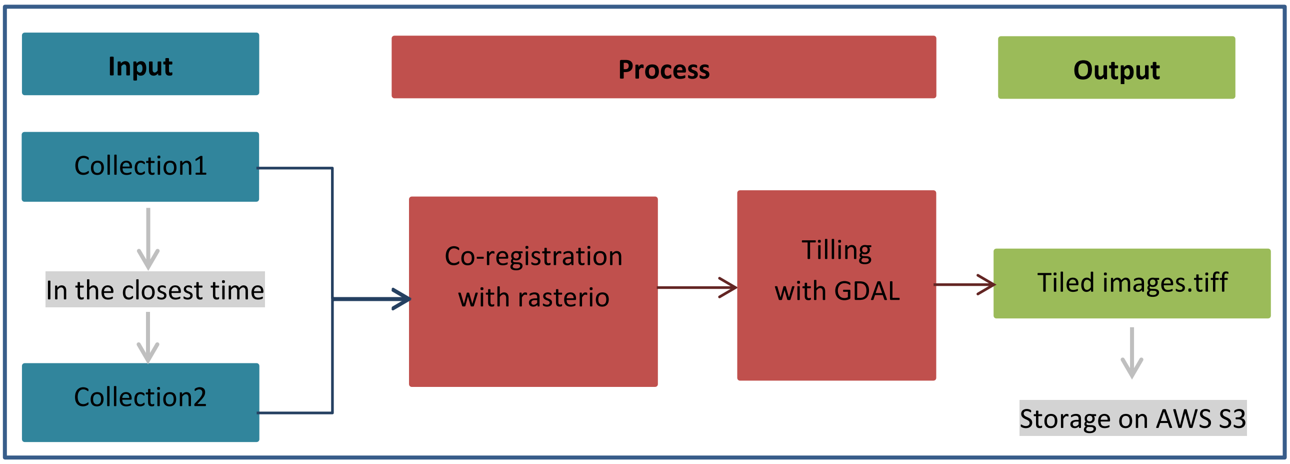
1. Sentinel1 images -> VV and VH bands
2. Sentinel2_L1C images -> True color
3. Sentinel2_L2A images -> True color
4. Sentinel2_L1C images -> Separated bands
5.



CH4 Figure 10: The workflow of request one type of data collection.

3.4.1.1 Request two types of data collection

This algorithm includes the same previous collection with the same characteristics of every two returned images from different collections which have been taken at the same time or in the available closest time.



CH 4 Figure 11: The workflow of request two types of data collection.

3.4.2 Data Preparation

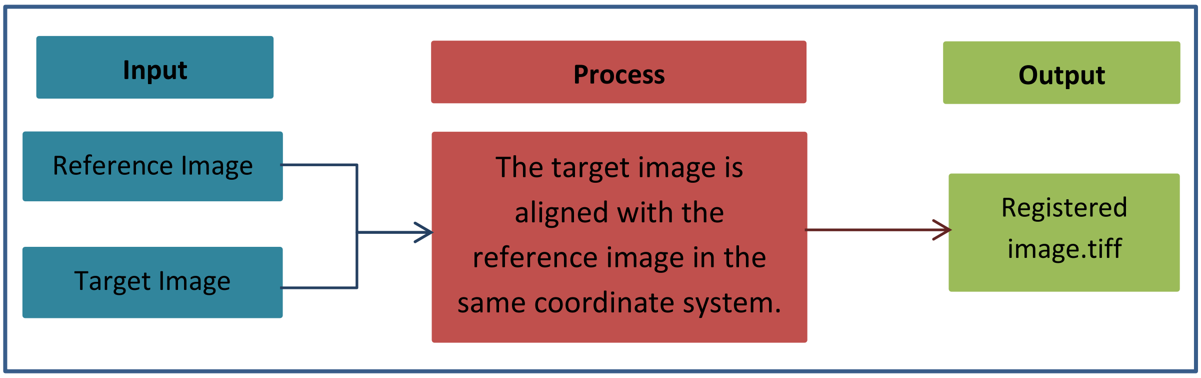
3.4.2.1 Data Co-registration

The quality of remotely sensed data can be negatively impacted by a range of factors, such as atmospheric conditions, the rotation of the Earth, the motion of satellites or aircraft, and the curvature of the Earth. To address these issues, Image registration, the process of transforming multiple sets of images or spatially referenced data into a single coordinate system, is used. If image data is acquired at different times or from different sources, they may be registered in different coordinate systems, which must be co-registered into a single coordinate system for subsequent analyses **[26].** Accurate registration is crucial for change detection techniques, which may produce false differences if misaligned. To ensure maximum spatial fidelity, it is recommended to co-register all data and geo-reference it using ground control points **[1].**

In order to automize **QGIS Co-registration plugin, which requires manual effort to make the images registered,** the algorithm uses the python QGIS plugin Repository, the following steps are executed by the algorithm:

1. Converting the target image using a pixel alignment process based on a reference image and then generating a new raster file based on the target image with all properties from the reference **[27].**
2. Providing automatically Co-registration of downloaded data using rasterio library, a python library allows access to geospatial raster data **[24]**, and gdal library, a translator library for raster and vector geospatial data formats **[25]**.

From the request of two types of data collection, two types of images are provided; one will be used as a reference image and the other will be used as a target image based on user request as the workflow in figure 12 shown.



CH4 Figure 12: The workflow of data co-registration

3.4.2.2 Data Tilling

Data Tiling is a process to split a large area into smaller ones **[28]**.

Based on the user input, the algorithm patches a registered image into smaller ones with each having width and height predetermined in the request. Later tilling the data technique will be shown.

3.5 The user interface

The user has the right to choose some interesting parameters for the images that will be downloaded.

These parameters are:

1. Area of interest as a polygon shape
2. Type of available data collection provided in the algorithm.
3. Time range of the returned images; the start time and the end time
4. Type of available bands of data collection provided in the algorithm.
5. Number of the returned images
6. Number of maximum cloud coverage

To prepare the request, the request should be sent in JSON file format to an input bucket on AWS and the images will be downloaded automatically on an output bucket.

So firstly, what is the JSON file? And what is the body Schema to send the request?

JSON, abbreviated as JavaScript Object Notation, is a syntax for defining data interchange formats that is both lightweight and text based. It is independent of any particular programming language, although it originated from ECMAScript. JSON provides a concise set of structuring rules that enable the representation of structured data in a portable format. **[19]**

3.5.1 Body Schema for JSON file

There are two schemas depending on the user request: one type of data collection or two types of data collection:

1. For request one type of data collection


```
{"type": "Feature",
 "geometry": {
       "type": "Polygon",
       "coordinates": [………]
},
 "properties": {
  "task": "………",
  "Start_date": "………",
  "End_date": "………",
  "limit": ………,
  "band": ………,
  "maxcc": ………,
  "analysis": {
         "type": "None",
         "object": "None"
  }}}
```


CH4 Figure 13: JSON file schema for one type of data collection

2. For request two types of data collection


```
{"type": "Feature",
 "geometry": {
           "type": "Polygon",
           "coordinates": [……………]
},
 "collection": [
 {
  "task1": "………",
  "Start_date": "………",
  "End_date": "………",
  "limit": ………,
  "band": ………,
  "maxcc": ………
 },
 {
  "task2": "………",
  "Start_date": "dependent",
  "End_date": "task1",
  "limit":  "null",
  "band": ………,
  "maxcc": ………
 }
 ],
  "preparation": [
 {
   "task": "Coregistration",
   "reference": " task1",
   "target": " task2"
 },
 {
  "task": "Tiling",
  "width": ………,
  "height": ………
 } ],
  "analysis": {
             "type": "None",
             "object": "None"
 }}
```


4.Figure14: JSON file schema for two types of data collection CH

3.5.2 Summary of JSON file’s parameters

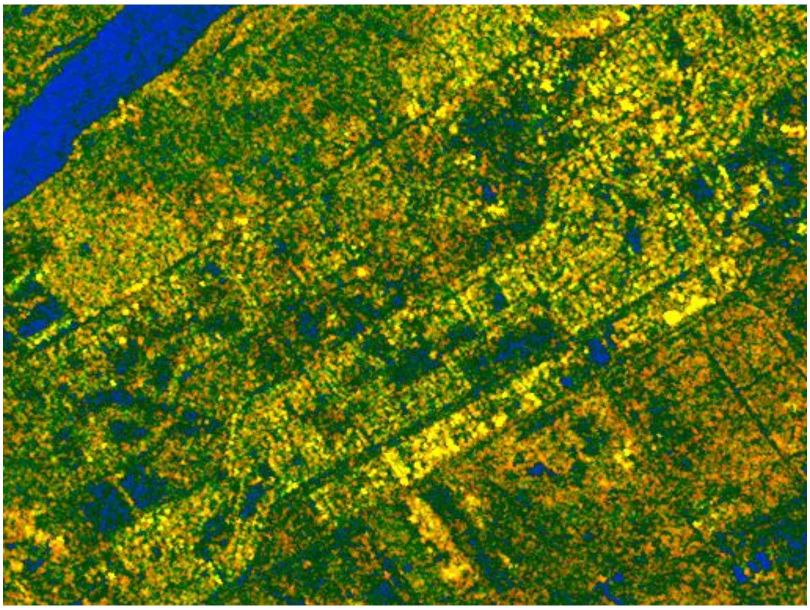
	coordinates	Task	start _date	end_date	band	limit	maxcc	width	height	Result
Meaning	Area of interest in polygon shape	Data collection name	Beginning a time interval	Ending a time interval	bands according to data collection	Images number	Maximum cloud coverage	Width of tiled image	height of tiled image	
Available data	[east Longitude, north latitude], [west Longitude, north latitude], [west Longitude, south latitude], [east Longitude, south latitude], [east Longitude, north latitude]	SENTINEL1 SENTINEL2 SENTINEL2_L1C SENTINEL2_L2A	year-month-day	year-month-day		0,1, 2,	0.0,0. 1,	64,128,512,	64,128,256,512,	
Example For request one type of data collection	[31.12604, 29.046599], [31.07268, 29.046599], [31.07268, 29.007827], [31.12604, 29.007827], [31.12604, 29.046599]	SENTINEL1 (VV+VH Polarization)	2023-04-01	2023-05-01	“”	1	0.0	----	----	Figure 15
		SENTINEL2_L1C	2023-04-01	2023-04-10	“True_color”	1	0.0	----	----	Figure 16
		SENTINEL2_L2A	2023-04-01	2023-04-10	“True_color”	2	0.0	----	----	Figure 17
		SENTINEL2	2023-04-01	2023-04-10	[“B01”, “B02”, “B05”]	1	0.1	----	----	Figure 18
Example For request two types of data collection	[31.12604, 29.046599], [31.07268, 29.046599], [31.07268, 29.007827], [31.12604, 29.007827], [31.12604, 29.046599]	SENTINEL1 SENTINEL2_L1C	2023-04-01	2023-04-02	“” “True_color”	1	0.0 0.0	128	128	Figure **22**
		SENTINEL2_L1C SENTINEL1	2023-04-01	2023-04-02	“True_color”	1	0.0 0.0	256	256	Figure 26

CH4 Table 6 : Summary of JSON file’s parameters

4. Results

After creating JSON file as previously explained, the results here are the results of running the JSON file on the server which already contains the algorithm. Each result here responds to examples in the JSON file's parameter in table6.

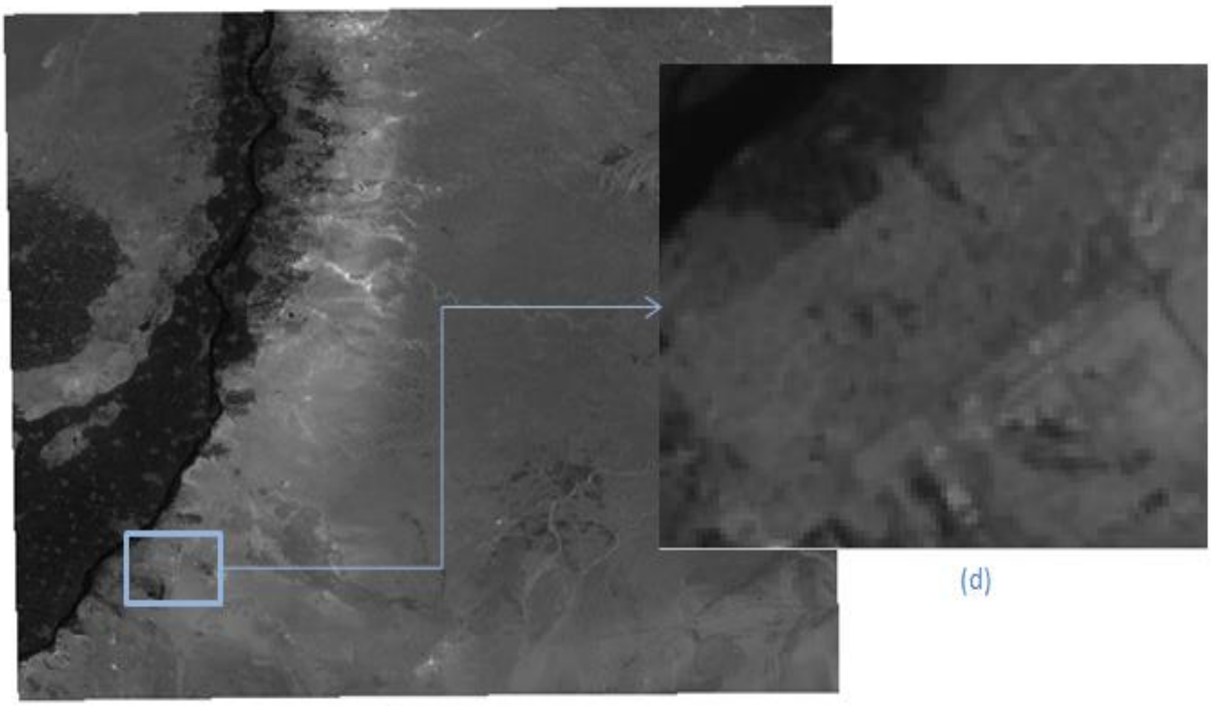
4.1 One type of data collection



CH4 Figure15: SENTINEL1 Image taken on 2023-04-01 15:56:36

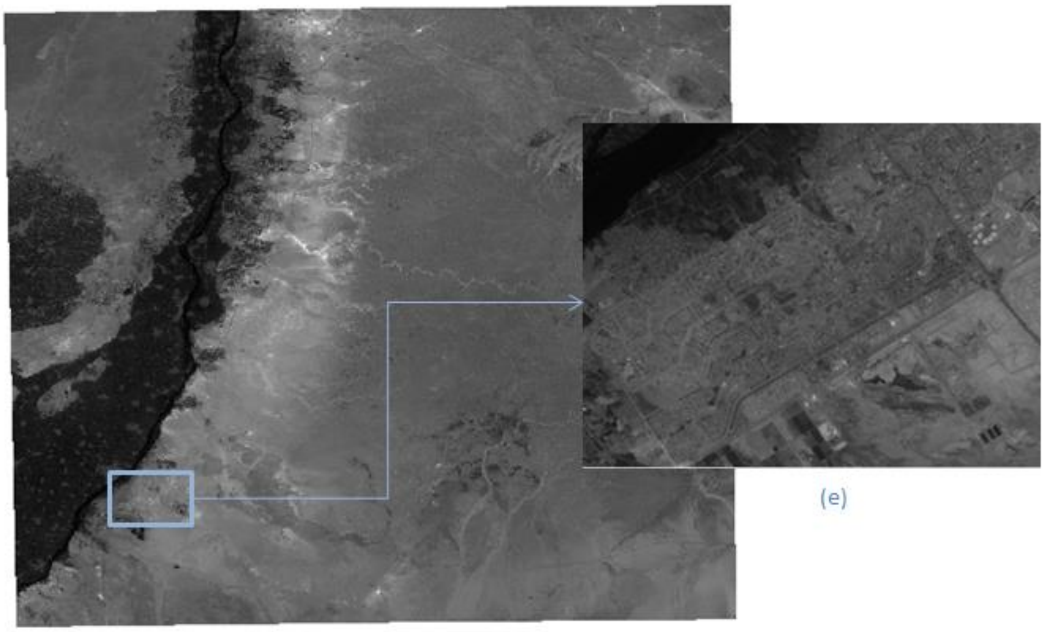
Figure 16: SENTINEL2_L1C Image taken on 2023-04-13 08:42:07
؛CH4

CH4. Figure 17: SENTINEL2_L2A Image taken on 2023-04-13 08:42:07



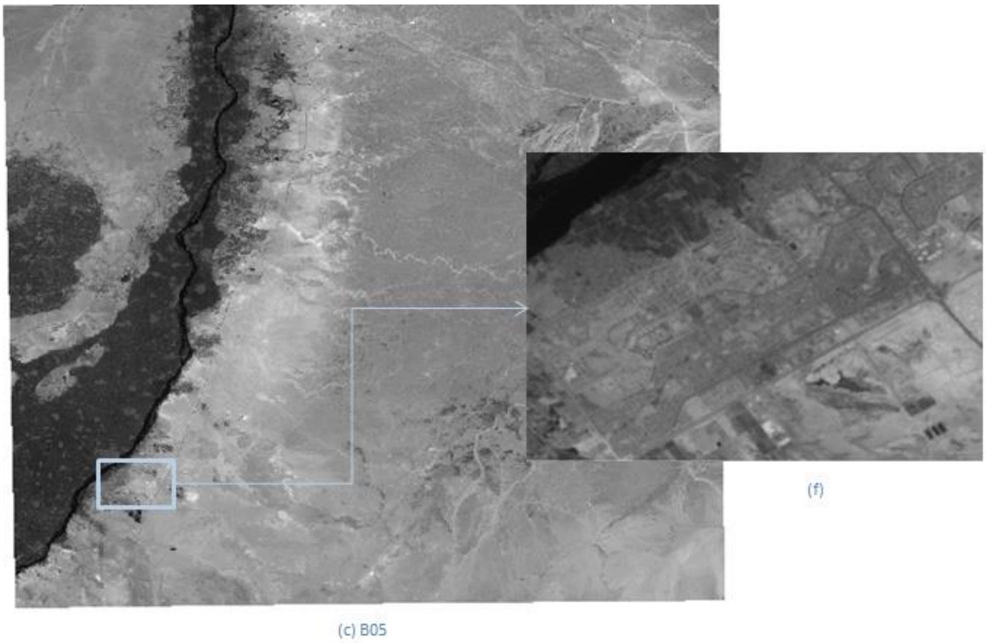
(d)

(a) B01



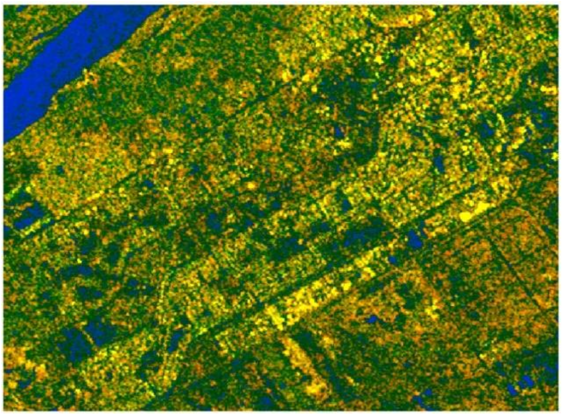
(e)

(b) B02



CH4 Figure 18: Sentinel2_L1C image separated bands taken on 2023-05-23: resolution of (a) B01, (b) B02 and (c) B05 are 60m, 10m and 20m respectively. And (d), (e) and (f) images are small areas maximized from the downloaded bands.

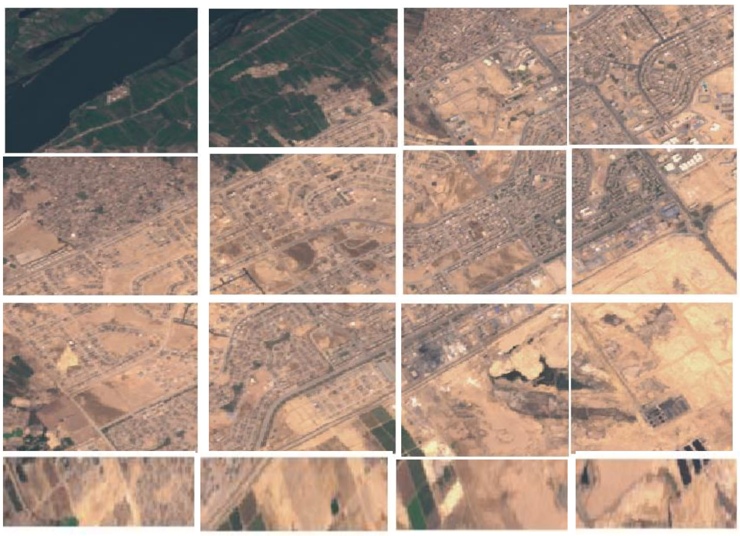
4.2 Two types of data collection



CH4 Figure 19: Sentinel1 Image

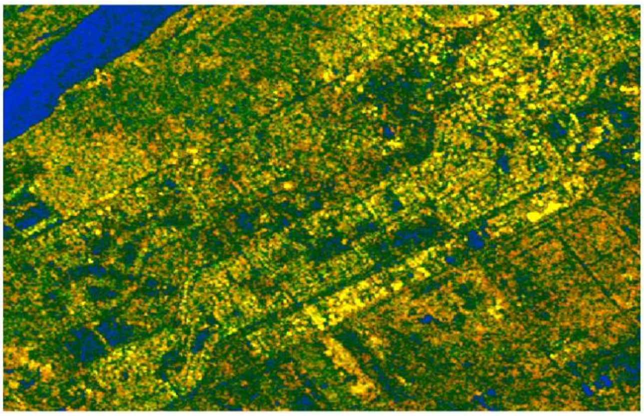
CH4 Figure 20: Sentinel2_L1C image.

∘CH4 Figure 21: Co-registered Sentinel2_L1C image with Sentinel1 as reference

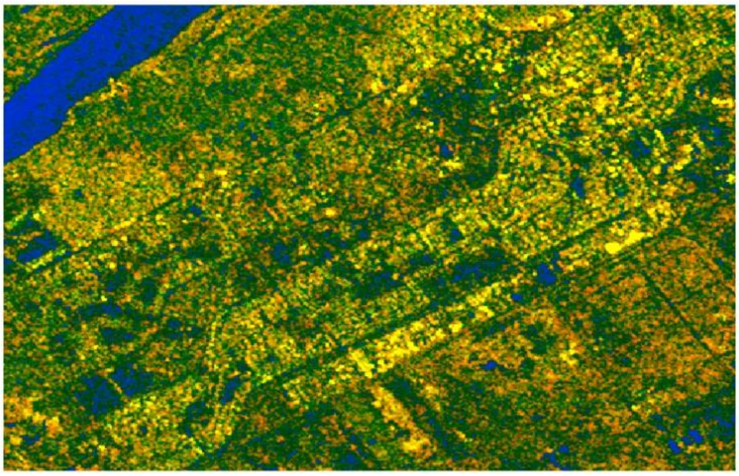


CH4 Figure 22: Tiles (128X128) of Co-registered Sentinel2_L1C

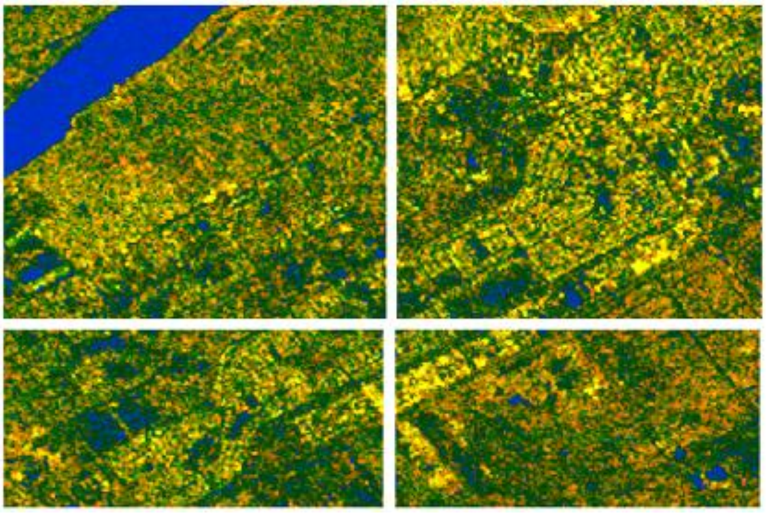
CH4 Figure23: Sentinel2_LIC image.



CH4 Figure 24: Sentinel Image



CH4 Figure 25:CO-registed Sentinel1 image with Sentinel12_LIC as reference 1



CH4 Figure 26: Tiles (256x256) of Co-registered Sentinel1 image.

5.Conclusion

In conclusion, this thesis has presented a comprehensive workflow for gathering and post-processing satellite imagery datasets. The workflow includes various stages, such as data acquisition, pre-processing, co-registration, and patching. Through the implementation of this workflow, it is possible to automate many of the time-consuming and resource-intensive tasks associated with satellite imagery processing, enabling faster and more efficient analysis of the data. The workflow has been demonstrated on a range of satellite imagery datasets, and the results have shown the effectiveness and accuracy of the automated approach.

Overall, this thesis has demonstrated the importance of satellite imagery datasets in various fields and the potential benefits of automating the post-processing of these datasets. By providing a detailed workflow, this thesis aimed to contribute to the development of more efficient and effective methods for analyzing satellite imagery data. It is hoped that this thesis will inspire further research in this area and help to advance the field of remote sensing, enabling us to better understand and monitor our planet's changing landscape.

5.CROP CLASSIFICATION

Abstract

Agriculture has always thrived on innovation, changing as the world changes. Today's agriculture requires a different kind of innovation to meet the demands of a growing, increasingly urban population, and the pressures of a changing climate. These factors place importance on an integrated approach to regenerative agriculture that optimizes the use of natural resources, farm inputs, and farmers' return on investment. Agricultural monitoring plays a crucial role in ensuring food security, optimizing resource management, and understanding the dynamics of crop production. The need for accurate and up-to-date crop type maps is paramount. These maps provide valuable information about the spatial distribution and extent of specific crops, such as rice, almond, and corn, enabling better decision-making and planning for agricultural activities. The availability of improved operational global agricultural monitoring systems offers significant benefits to stakeholders involved in agriculture and land management. These maps provide valuable insights into the spatial distribution of crops, allowing policymakers, farmers, and researchers to optimize resource allocation, monitor crop health, detect anomalies or pest outbreaks, and assess the impact of climate change on agricultural productivity. All season, every season, in every country around the world, farmers collect and evaluate large amounts of data. From seed varieties planted to inputs applied to harvested to the success of operational practices, farmer increasingly rely on data science and the advent of digital technologies as well as new opportunities to use that data to improve operations and outcomes. Our farmers and scientists consider the complex interactions between the growing environment, seed genetics and farm management practices in their quest to unlock the future of prescriptive agriculture. In the future, leveraging agronomic modeling techniques in partnership with farmers will drive the adoption of the right products, applied at the right times, and in the right ways to optimize farm sustainability and productivity. We believe that supporting farmers around the world as they adopt digital tools - and the advanced insights that come from them - is the key to farming's future. Next generation agricultural solutions, where leading seed genetics are paired with crop protection products with unparalleled safety and an underlying platform in our map's suite of digital tools, is what will unlock the sustainable farm productivity needed to feed a changing world.

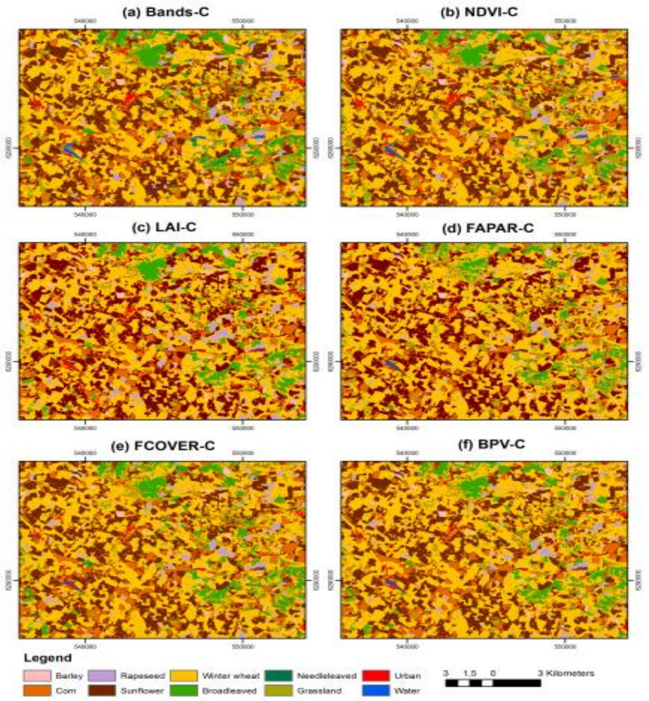
1. Introduction

With ongoing global population growth and the corresponding rise in demand for food, feed, fiber, and fuel, agriculture takes a key role to sustain future livelihood. Competition for land within the agricultural sector, but also with other land uses, can lead to agricultural intensification and expansion, often with negative impacts on ecosystems, like deforestation and biodiversity. The Global Rice map or any other field is an interactive tool that showcases the worldwide distribution of rice, a crucial source of nutrition. It plays a vital role in forecasting, monitoring, and assessing rice production at local, regional, and national levels. This allows us to evaluate the efficient use of limited resources such as land, water, and fertilizers, understand the impact of climate on rice cultivation, and address food security concerns. [1] Rice is one of the major crops on earth, feeding more than half of the world's population, and its cultivation area also exceeds 12 % of the world's cultivated land, playing a crucial function in the food security of human beings. Moreover, rice planting has far-reaching effects on water consumption and global warming. Therefore, early access to the spatial distribution of rice is critical for agricultural systems and food security. This search aims to explore the advancements made in operational global agricultural monitoring techniques specifically focused on rice, almond, and corn crops in California and USA. The utilization of satellite imagery and remote sensing data offers an efficient and reliable means of capturing essential agricultural information over large areas. By harnessing these technologies, researchers and organizations have made significant strides in improving the accuracy and timeliness of crop type mapping, facilitating better agricultural management practices, and supporting policy formulation. The main goal of the crop type classification is to create a map/image where each point on it will have a class of what is growing there. Numerous studies have confirmed that remote sensing (RS) technology provides an accurate, fast, and low-consumption technical approach for paddy rice mapping (PRM). The RS images used for PRM generally include optical images (OIs) and SAR images (SARIs). Different types of vegetation indices (VIs) constructed using spectral information, such as the normalized difference vegetation index (NDVI), The Normalized Difference Water Index (NDWI) is used to highlight open water features in a satellite image, allowing a water body to "stand out" against the soil and vegetation.

2. Literature Review

Paper [**1**] is about the use of biophysical variables and remote sensing data for land cover and crop type classification. It explains the importance of these variables in capturing vegetation dynamics and improving classification accuracy. The document discusses the extraction and integration of high-resolution multi-sensor biophysical time series and compares their performance with spectral information and NDVI. It also explores the impact of different acquisition dates on classification accuracy and the optimal length of time series for accurate classification. The document provides insights into the potential of using biophysical variables for land cover mapping and crop monitoring. Overall, it highlights the importance of temporal trajectory and the potential benefits of using biophysical variables in classification tasks. This article investigates the potential of structural biophysical variables as common parameters to consistently combine multi-sensor time series and to exploit them for land/crop cover classification. Artificial neural networks were trained based on a radiative transfer model in order to retrieve high resolution LAI, FAPAR and FCOVER from Landsat-8 and SPOT-4. The correlation coefficients between field measurements and the retrieved biophysical variables were 0.83, 0.85 and 0.79 for LAI, FAPAR and FCOVER, respectively. The retrieved biophysical variables' time series displayed consistent average temporal trajectories, even though the class variability and signal-to-noise ratio increased compared to NDVI. Six random forest classifiers were trained and applied along the season with different inputs: spectral bands, NDVI, as well as FAPAR, LAI and FCOVER, separately and jointly. Classifications with structural biophysical variables reached end-of-season overall accuracies ranging from 73%–76% when used alone and 77% when used jointly. This corresponds to 90% and 95% of the accuracy level achieved with the spectral bands and NDVI. FCOVER appears to be the most promising biophysical variable for classification. When assuming that the cropland extent is known, crop type classification reaches 89% with spectral information, 87%

with the NDVI and 81%–84% with biophysical variables. Six random forest classifiers were trained along the season with the spectral bands (Bands-C), NDVI (NDVI-C) and three biophysical variables separately (LAI-C, FAPAR-C and FCOVER-C) and jointly (BPV-C). End of season maps show akin general patterns for the six classifiers. Nevertheless, a salt and pepper effect are more visible on maps derived from biophysical variables than on the NDVI-C and Bands-C maps (Figure 29).



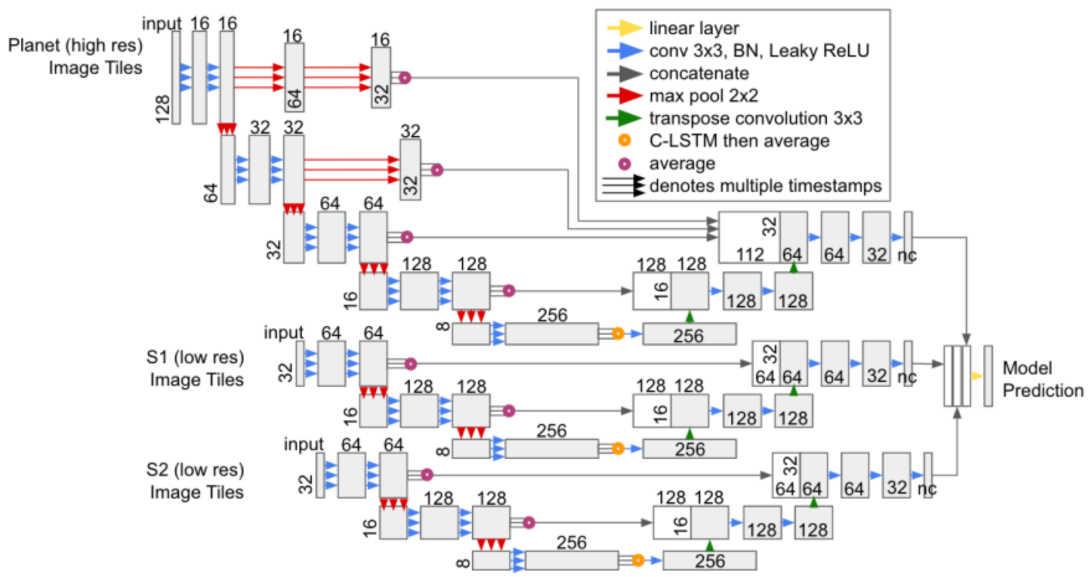
CH 5 Figure 1: End-of-season maps for the six classifiers: (a) Band-C; (b) NDVI-C; (c) LAI-C; (d) FAPAR-C; (e) FCOVER-C; (f) BPV-C. Bands-C provides the most accurate classification accuracy, followed by NDVI-C and the biophysical variables. The pepper and salt are more visible on maps derived from biophysical variables than on the NDVI-C and Bands-C maps. This ought to be related to the lower SNR observed for biophysical variables.

In Paper [**2**], multi-temporal Sentinel-2 data acquired in the growing season in 2019 were applied to the random forest algorithm to generate the crop classification map at 10 m spatial resolution for the Shiyang River Basin. Four experiments with different combinations of feature sets were carried out to explore which Sentinel-2 information was more effective for higher crop classification accuracy. The results showed that the augment of multi-spectral and multi-temporal information of Sentinel-2 improved the accuracy of crop classification remarkably, and the improvement was firmly related to strategies of feature selections. Compared with other bands, red-edge band 1 (RE-1) and shortwave-infrared band 1 (SWIR-1) of Sentinel-2 showed a higher competence in crop classification. The combined application of images in the early, middle, and late crop growth stage is significant for achieving optimal performance. A relatively accurate classification (overall accuracy = 0.94) was obtained by utilizing the pivotal spectral bands and dates of image. In addition, a crop map with a satisfied accuracy (overall accuracy > 0.9) could be generated as early as late July. This study gave an inspiration in selecting targeted spectral bands and period of images for acquiring more accurate and timelier crop map. The proposed method could be transferred to other arid areas with similar agriculture structure and crop phenology. In this study, they chose the RF model [4] for crop classification. The RF algorithm is an efficient algorithm based on an ensemble idea proposed by Breiman that consists of multiple decision trees or classified regression trees. Random forest algorithm can effectively reduce model overfitting by introducing randomness of training samples and classification features. Several subsamples are extracted from training samples by a random sampling method which is a bootstrapping method. The scikit–learn package RandomForestClassifier in Python was used in our work to implement the RF algorithm [5]. Two predominant parameters determine the performance of the algorithm. One is the number of decision trees. Previous studies suggested the classification error or overall accuracy converges with the increase of the number of trees [6,7]. they tested the value of 100, 300, 500 and 700 and found that 700 did not significantly improve the accuracy. Taking into account the computing time, they finally selected 500 as the number of decision trees to permit the convergence of the out-of-bag error. Another parameter is the number of features involved in the training of each decision tree. It was set to the squared root of number of the input features as lots of literature recommended [8]. With multi-temporal information, the overall accuracy increased to over 0.94 for all the spectral combinations and the highest overall accuracy was 0.95 as showed in table [1].

Combinations	Band of Sentinel-2	Overall Accuracy
C2Bs	RE-1, SWIR-1	0.94
C3Bs	RE-1, NIR, SWIR-1	0.95
C4Bs	GREEN, RE-1, NIR, SWIR-1	0.95
C5Bs	BLUE, Red, RE-1, NIR, SWIR-1	0.95
C6Bs	BLUE, GREEN, Red, RE-1, NIR, SWIR-1	0.95
C7Bs	All bands without Red and SWIR-2	0.95
C8Bs	All bands without Red	0.95
C9Bs	All bands	0.95

CH 5 Table 1: Best combinations for different groups when using spectral combinations with full time series.

Paper [3], They provide the first crop type semantic segmentation dataset of small holder farms, specifically in Ghana and South Sudan. They are also the first to utilize high resolution, high frequency satellite data in segmenting small holder farms. Despite the challenges, we achieve an average F1 score and overall accuracy of 57.3 and 60.9% in Ghana and 69.7 and 85.3% in South Sudan. Additionally, our approach outperforms the state-of-the-art method in a data-rich setting of Germany by over 8 points in F1 and 6 points in accuracy. They compare performance between a 3D U-Net and a model that incorporates both CNNs and RNNs for semantic segmentation of multi-temporal, multi-spatial satellite images. To gain further insight into the sequence models and contributing attributes, they explore ablation studies and compare with a random forest baseline. They predict crop type with reasonable performance in Ghana and South Sudan where data is limited and of poor quality due to high cloud cover, class imbalance, and lack of labels. When applied on a large dataset in Germany, they surpass state-ofthe-art performance on this task. They release the full datasets and code repository and hope to encourage the development of crop type segmentation systems for small holder farms.



CH 5 Figure 2: The 2D U-Net + CLSTM model architecture used in this study; "nc" denotes number of output classes.

3. Methodology

3.1 Dataset and Features.

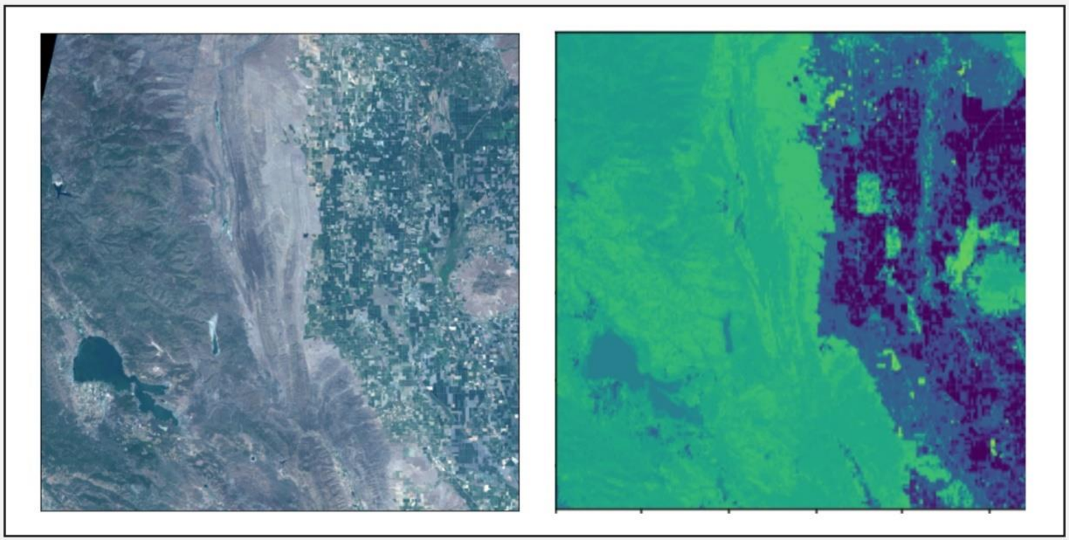
Locations and labels: Our dataset are made up of sparse ground truth labels of crop fields in USA, California, and Egypt. Ground truth labels consist of geo- referenced polygons, where each polygon represents an agricultural field boundary with a crop type label. we use Labels data from **CropScape** for the 2021 growing season. The most field in California is Rice and the most field in USA is corn So We build two Models one use to predict Rice and other to predict corn.

```
# print the first 5 crope in image
print('The First crope is Rice = ', df['count'].nlargest(5)[0:1])
print('The second crope is Almonds = ', df['count'].nlargest(5)[1:2])
print('The third  crope is Fallow/Idle Cropland = ', df['count'].nlargest(5)[2:3])
print('The Fourth crope is  Shrubland = ', df['count'].nlargest(5)[3:4])
print('The fifth crope is Walnuts = ', df['count'].nlargest(5)[4:5])
```

```
The First crope is Rice =  3    2202745
Name: count, dtype: int64
The second crope is Almonds =  75    1461436
Name: count, dtype: int64
The third  crope is Fallow/Idle Cropland =  61    1131661
Name: count, dtype: int64
The Fourth crope is  Shrubland =  152    864309
Name: count, dtype: int64
The fifth crope is Walnuts =  76    829224
Name: count, dtype: int64
```

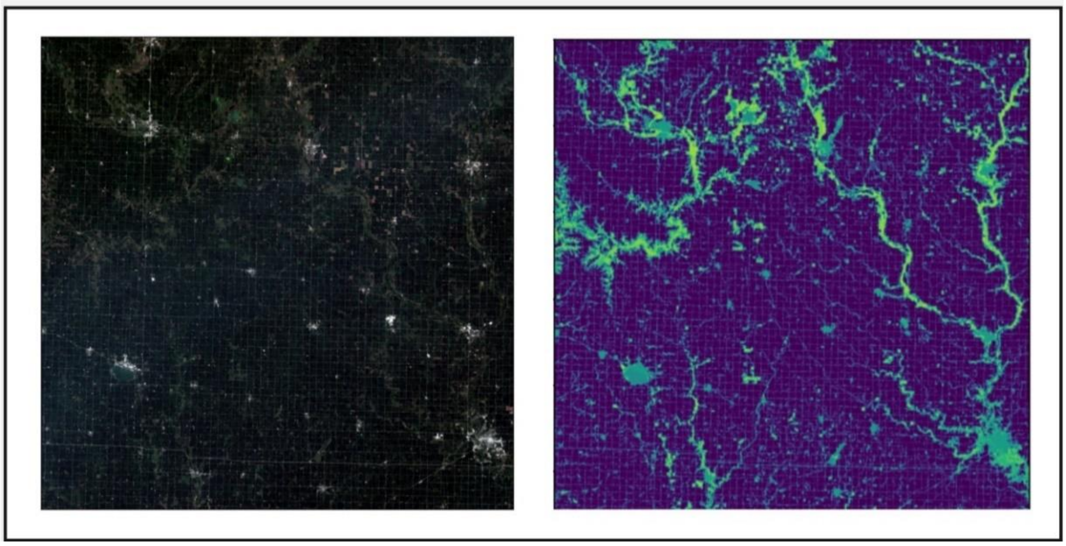
CH 5 Figure 3 the first 5 crops in California in 2021.

As we shown in figure 3 the first 5 crops in California in 2021. In California, focus on rice, we replace all crops as background.



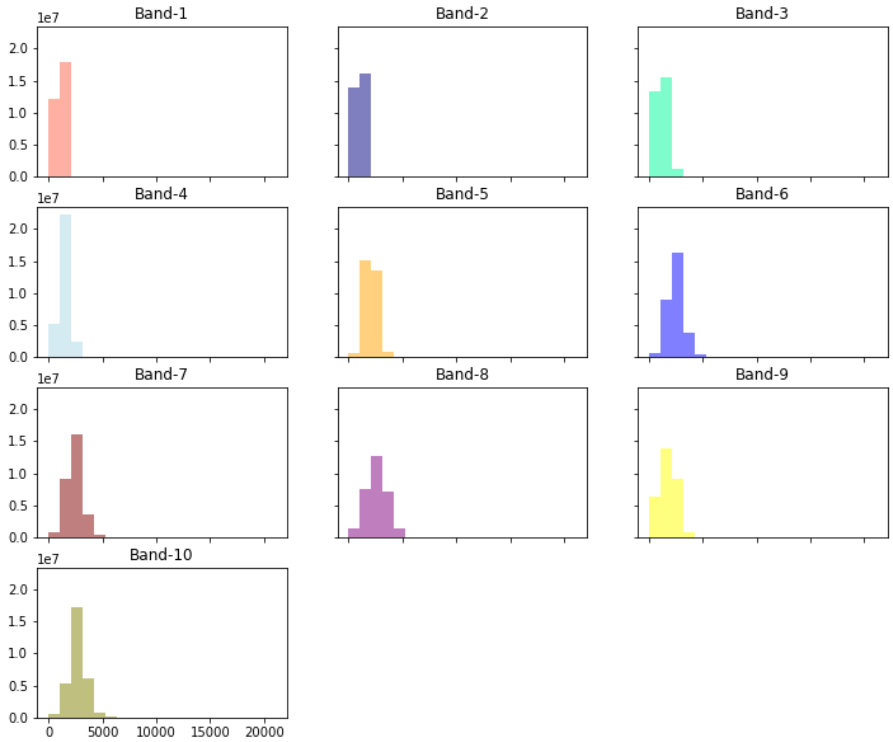
CH 5 Figure 4: Sentine-l2 rgb & ref image in califorina.

In the USA, we focus on corn; we replace all crops as background.

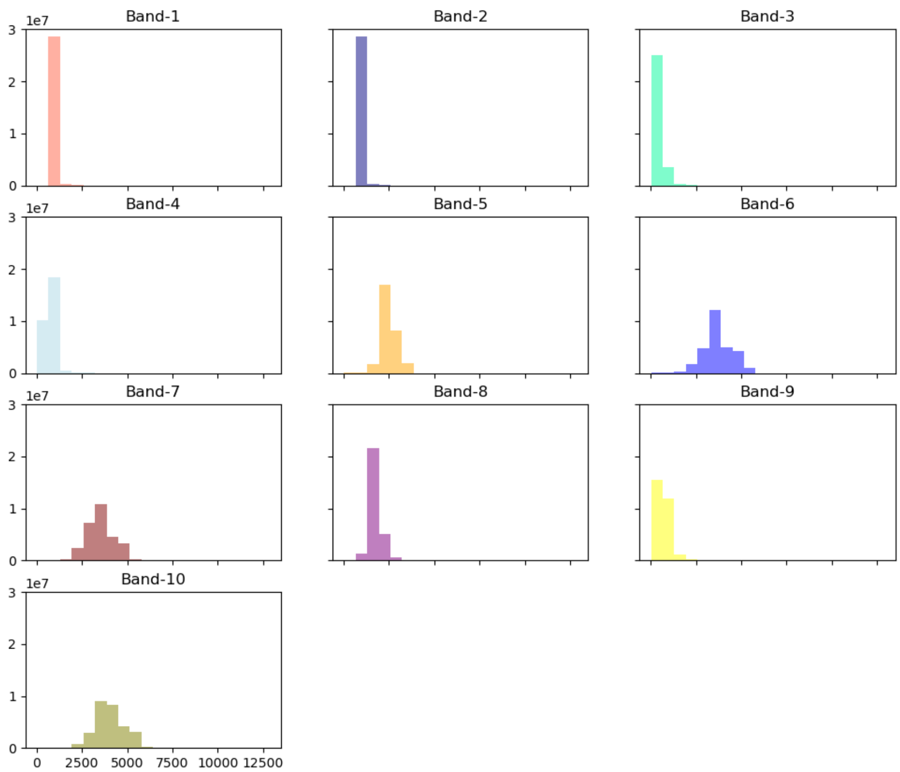


CH 5 Figure 5: Sentine-l2 rgb & ref image in USA.

Input features: We create inputs to the model by mapping S1 and S2 to the labeled locations within our dataset. As input features, we use ten S2 bands (all bands except 1, 9 and 10), and three S1 bands. Both Sentinel satellites have a 20m spatial resolution. With high cloud cover and small field sizes, we believe incorporating Planet imagery will be beneficial. In experiments we also use Normalized Difference Vegetation Index (NDVI), Normalized Difference Vegetation Index (NDWI) and Normalized Difference Moisture Index (NDMI).



CH 5 Figure 6: Bands Rice Signiture.



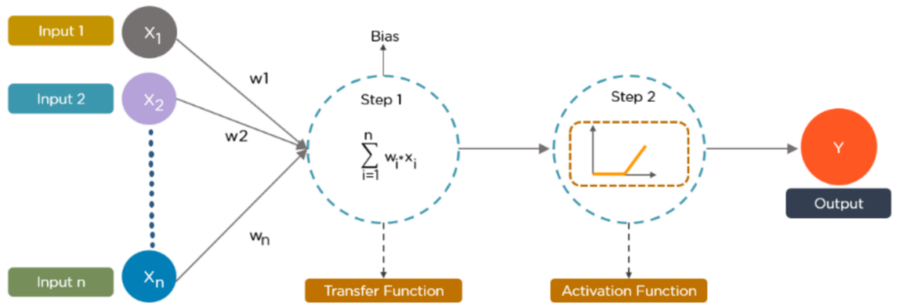
CH 5 Figure 7 Bands Corn Signiture.

3.2 Methods.

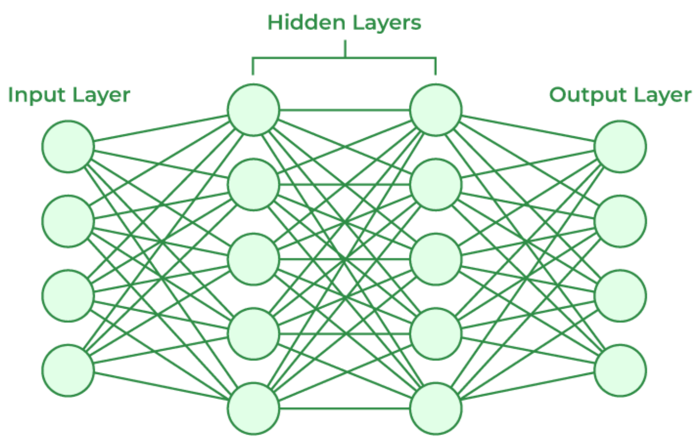
After downloading the S2 & S1, we co-register S1 with S2 and co-register label with S2 as both have different resolutions. Next step we resample all bands into same resolution 20m. Next step we combine all bands and merge S2&S1. Next step we tile the images into small patches to be prepared for the AI Model (128,128). Now, Dataset should be (1764, 128, 128, 13). 1764 this number of Batches, and x, y (128,128), and 13 is the number of bands. Now, the Dataset is prepared. So, we will build models. But Before we Talk about U-Net we talk about some definition.

3.2.1 Neural Network.

An artificial neuron network is a computational model that mimics the way nerve cells work in the human brain. There are more complicated and high-end models in the DL approach. However, ANN is a vital element in all the models in DL.



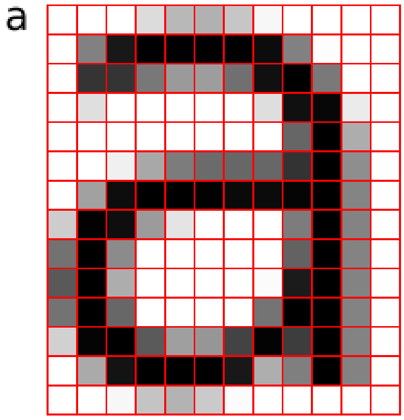
CH 5 Figure 8: neural network



CH 5 Figure 9: neural network with hidden layers

3.2.2. Convolutional Neural Networks

A Convolutional Neural Network, also known as CNN or Conv-Net, is a class of neural networks that specializes in processing data that has a grid-like topology, such as an image. A digital image is a binary representation of visual data. It contains a series of pixels arranged in a grid-like fashion that contains pixel values to denote how bright and what color each pixel should be.

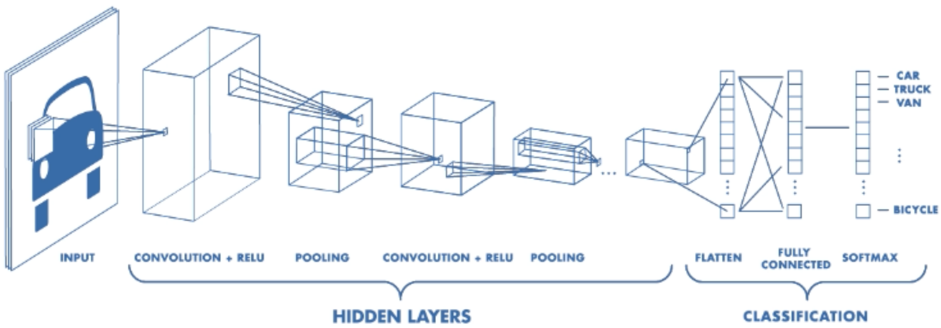


1.0	1.0	1.0	0.9	0.6	0.6	0.6	1.0	1.0	1.0	1.0	1.0
1.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0	1.0	1.0
1.0	0.2	0.2	0.5	0.6	0.6	0.5	0.0	0.0	0.5	1.0	1.0
1.0	0.9	1.0	1.0	1.0	1.0	1.0	0.9	0.0	0.0	0.9	1.0
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
1.0	1.0	1.0	0.5	0.5	0.5	0.5	0.5	0.4	0.0	0.5	1.0
1.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	1.0
0.9	0.0	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.6	1.0	1.0	1.0	1.0	1.0	0.5	0.0	0.5	1.0
0.5	0.0	0.7	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.5	1.0
0.6	0.0	0.6	1.0	1.0	1.0	1.0	0.5	0.0	0.0	0.5	1.0
0.9	0.1	0.0	0.6	0.7	0.7	0.5	0.0	0.5	0.0	0.5	1.0
1.0	0.7	0.1	0.0	0.0	0.0	0.1	0.9	0.8	0.0	0.5	1.0
1.0	1.0	1.0	0.8	0.8	0.9	1.0	1.0	1.0	1.0	1.0	1.0

CH 5 Figure 10: Convolutional Neural Networks

3.2.3 Convolutional Neural Network Architecture

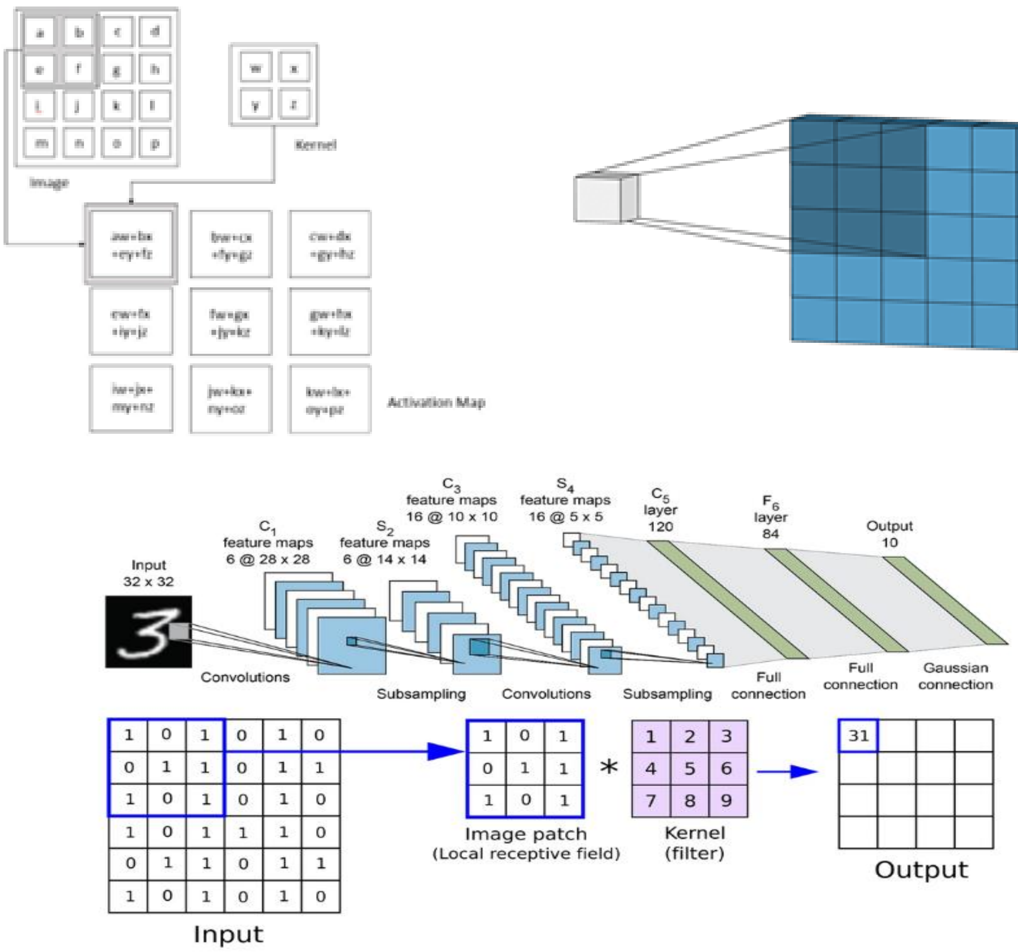
CNN typically has three layers: a convolutional layer, a pooling layer, and a fully connected layer.



CH 5 Figure 11: Convolutional Neural Network Architecture

3.2.4 Convolution Layer

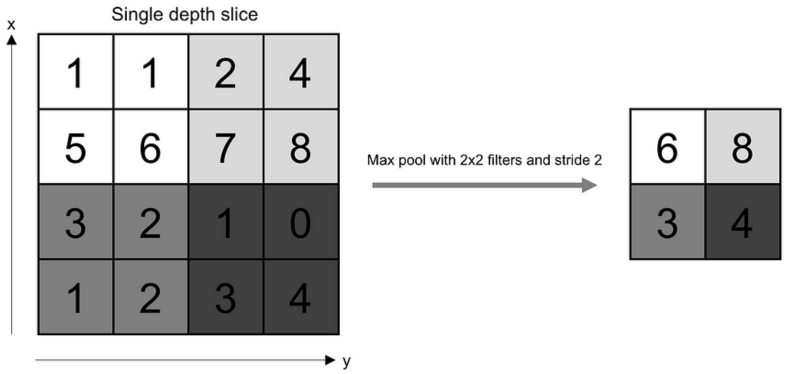
The convolution layer is the core building block of CNN. It carries the main portion of the network's computational load. This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.



CH 5 Figure 12: Convolutional layer

3.2.5 Pooling Layer

The pooling layer replaces the output of the network at certain locations by deriving a summary statistic of the nearby outputs. This helps in reducing the spatial size of the representation, which decreases the required amount of computation and weights. The pooling operation is processed on every slice of the representation individually. There are several pooling functions such as the average of the rectangular neighborhood, L2 norm of the rectangular neighborhood, and a weighted average based on the distance from the central pixel. However, the most popular process is max pooling, which reports the maximum output from the neighborhood.



CH 5 Figure 13: Pooling Operation.

3.2.6 Fully Connected Layer.

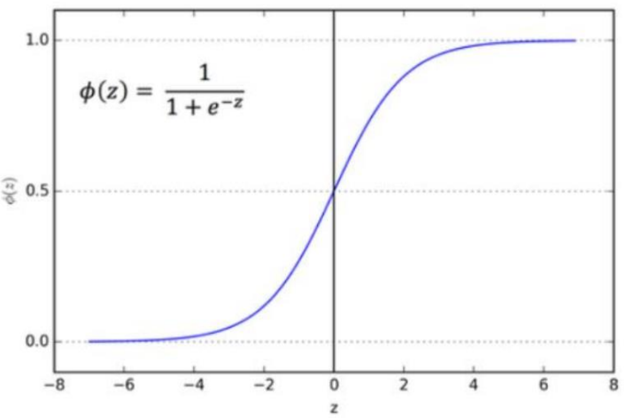
Neurons in this layer have full connectivity with all neurons in the preceding and succeeding layer as seen in regular FCNN. This is why it can be computed as usual by a matrix multiplication followed by a bias effect. The FC layer helps to map the representation between the input and the output.

3.2.7 Non-Linearity Layers

Since convolution is a linear operation and images are far from linear, non-linearity layers are often placed directly after the convolutional layer to introduce non-linearity to the activation map. There are several types of non-linear operations, the popular ones being:

3.2.7.1 Sigmoid

The sigmoid non-linearity has the mathematical form $\sigma(\kappa) = 1/(1+e^{-\kappa})$. It takes a real-valued number and "squashes" it into a range between 0 and 1. However, a very undesirable property of sigmoid is that when the activation is at either tail, the gradient becomes almost zero. If the local gradient becomes very small, then in backpropagation it will effectively "kill" the gradient. Also, if the data coming into the neuron is always positive, then the output of sigmoid will be either all positives or all negatives, resulting in a zig-zag dynamic of gradient updates for weight.

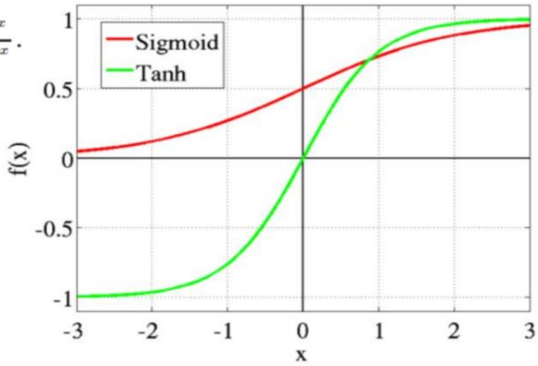


CH 5 Figure 14: Sigmoid Function.

3.2.7.2. Tanh

Tanh squashes a real-valued number to the range [-1, 1]. Like sigmoid, the activation saturates, but — unlike the sigmoid neurons — its output is zero centered.

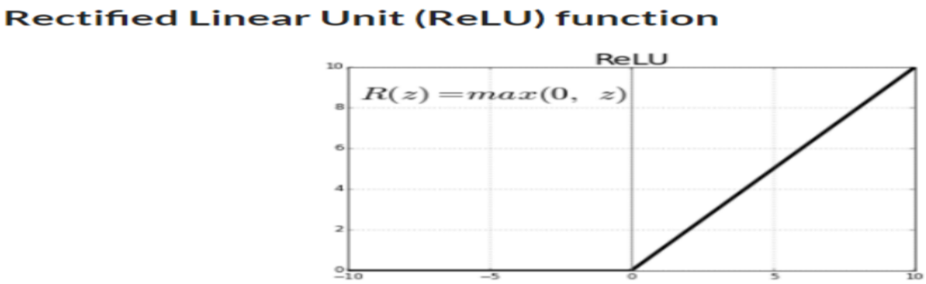
$$tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}.$$



CH 5 Figure 15: Sigmoid & Tanh.

3.2.7.3. ReLU

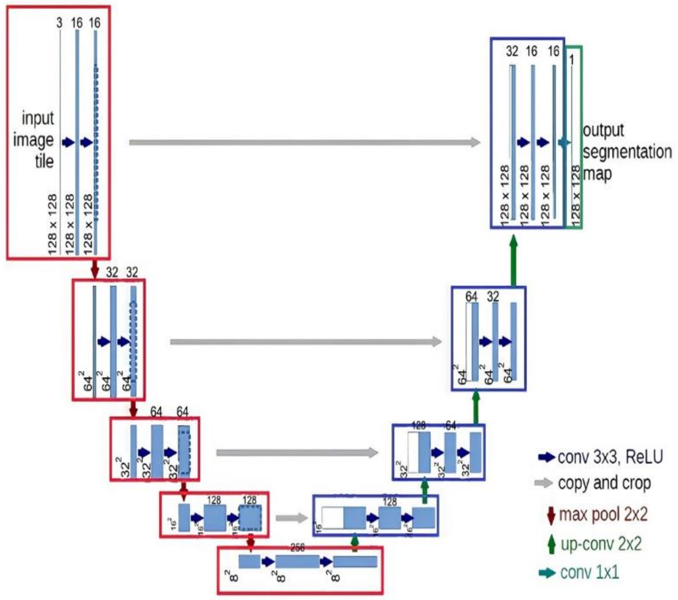
The Rectified Linear Unit (ReLU) has become very popular in the last few years. It computes the function $f(\kappa)$=max (0, κ). In other words, the activation is simply threshold at zero. In comparison to sigmoid and tanh, ReLU is more reliable and accelerates the convergence by six times. Unfortunately, a con is that ReLU can be fragile during training. A large gradient flowing through it can update it in such a way that the neuron will never get further updated. However, we can work on this by setting a [roper learning rate.



CH 5 Figure 16: Relu

In this work, we explore segmentation (2D U-Net). It evolved from the traditional convolutional neural network, was first designed, and applied in 2015 to process biomedical images. As a general convolutional neural network focuses its task on image classification, where input is an image and output are one label, but in this case, it requires us not only to distinguish whether there is a crop, but also to localize the area of this crop.

U-Net is dedicated to solving this problem. The reason it can localize and distinguish borders is by doing classification on every pixel, so the input and output share the same size.

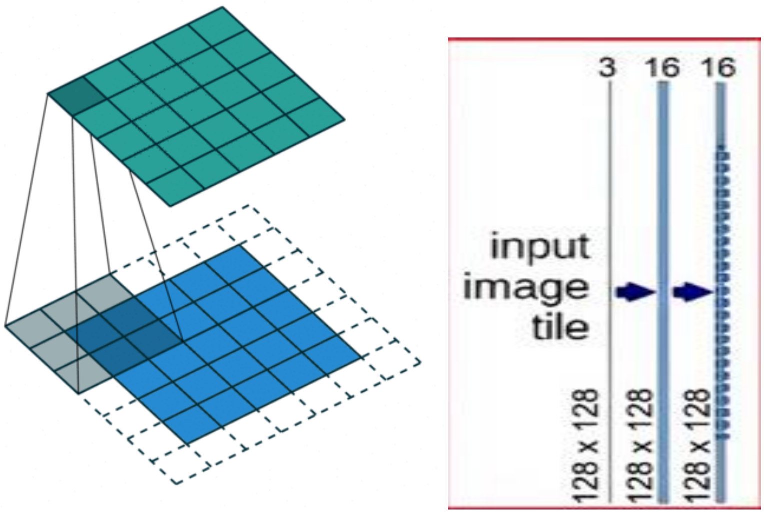


CH 5 Figure 17: U-Net architecture

3.2.8 Contracting Path

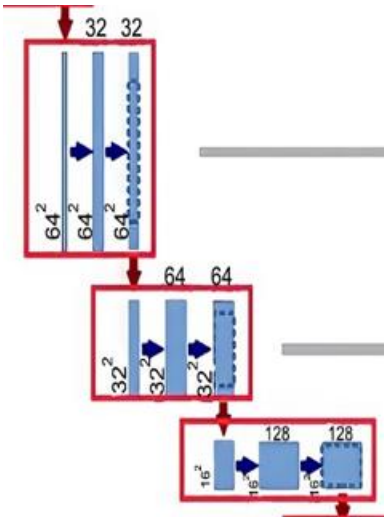
The contracting path follows the formula:

conv_layer1 -> conv_layer2 -> max_pooling -> dropout(optional)

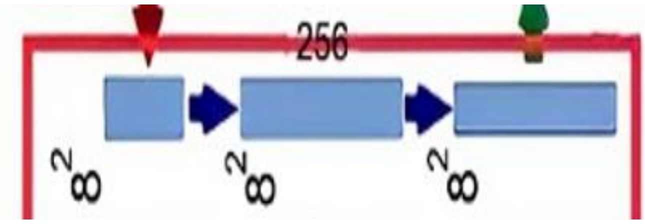


CH 5 Figure 18: two convolutional layers.

Notice that each process constitutes two convolutional layers, and the number of channel changes from 1 → 64, as convolution process will increase the depth of the image. The red arrow pointing down is the max pooling process which halves downsize of image (the size reduced from 128x128 → 64x64 is due to padding issues, but the implementation here uses padding= “same”). The process is repeated 3 more times:



CH 5 Figure 19: Contracting Path



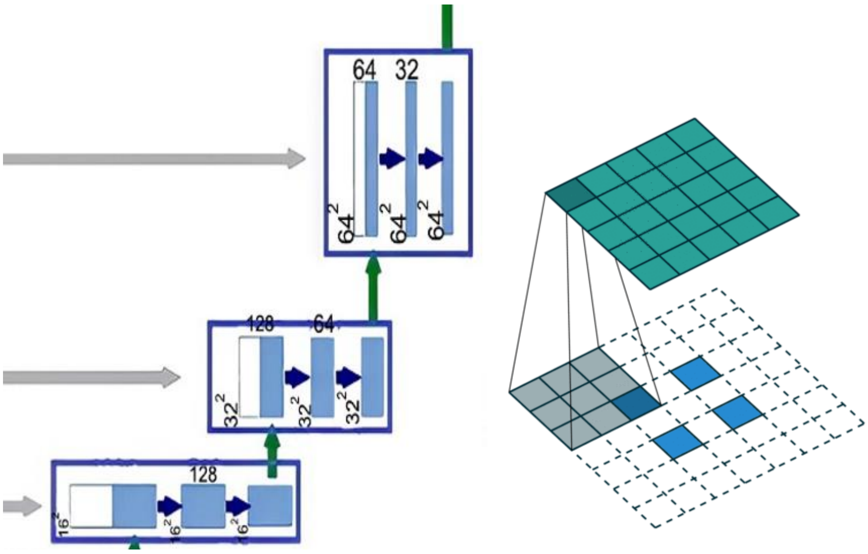
CH 5 Figure 20 bottle neck.

Now, we reach the bottle neck: The image at this moment has been resized to 8x8x256. Now let’s get to the expansive path.

3.2.9 Expansive Path

In the expansive path, the image is going to be upsized to its original size. The formula follows:

conv_2d_transpose -> concatenate -> conv_layer1 -> conv_layer2



CH 5 Figure 21: Expansive Path.

Transposed convolution is an up-sampling technic that expands the size of images. Basically, it does some padding on the original image followed by a convolution operation. After the transposed convolution, the image is upsized from 8x8x256 → 16x16x128, and then, this image is concatenated with the corresponding image from the contracting path and together makes an image of size 16x16x128. The reason here is to combine the information from the previous layers in order to get a more precise prediction. The last layer is a convolution layer with 1 filter of size 1x1.

3.2.10 Hyperparameters

Models were trained using a weighted combination of cross-entropy loss. Adam was used for parameter optimization. All the models were trained for 40 or 20 epochs depending on experiment to avoid over fitting. We applied dropout but did not modify it further in all experiments.

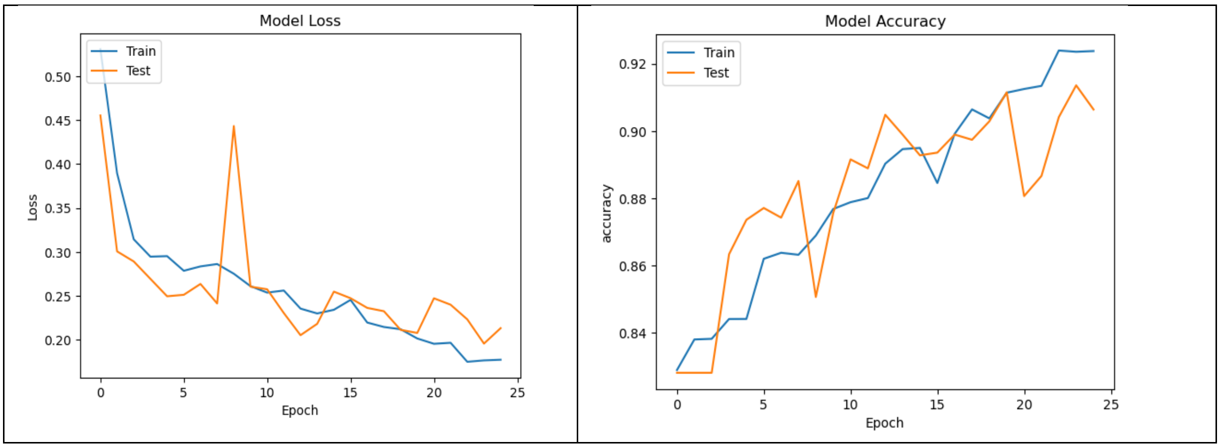
4. Analysis

4.1 Model 1

Models were trained using a weighted combination binary_crossentropy. Adam was used for optimization. All the models were trained for 25 epochs total. Start as 64 filter!

```python
#Contraction path
c1 = tf.keras.layers.Conv2D(16, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(s)
c1 = tf.keras.layers.Dropout(0.1)(c1)
c1 = tf.keras.layers.Conv2D(16, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c1)
p1 = tf.keras.layers.MaxPooling2D((2, 2))(c1)

c2 = tf.keras.layers.Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(p1)
c2 = tf.keras.layers.Dropout(0.1)(c2)
c2 = tf.keras.layers.Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c2)
p2 = tf.keras.layers.MaxPooling2D((2, 2))(c2)

c3 = tf.keras.layers.Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(p2)
c3 = tf.keras.layers.Dropout(0.2)(c3)
c3 = tf.keras.layers.Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c3)
p3 = tf.keras.layers.MaxPooling2D((2, 2))(c3)

c4 = tf.keras.layers.Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(p3)
c4 = tf.keras.layers.Dropout(0.2)(c4)
c4 = tf.keras.layers.Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c4)
p4 = tf.keras.layers.MaxPooling2D(pool_size=(2, 2))(c4)

#bottle neck
c5 = tf.keras.layers.Conv2D(256, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(p4)
c5 = tf.keras.layers.Dropout(0.3)(c5)
c5 = tf.keras.layers.Conv2D(256, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c5)

#Expansive path
u6 = tf.keras.layers.Conv2DTranspose(128, (2, 2), strides=(2, 2), padding='same')(c5)
u6 = tf.keras.layers.concatenate([u6, c4])
c6 = tf.keras.layers.Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(u6)
c6 = tf.keras.layers.Dropout(0.2)(c6)
c6 = tf.keras.layers.Conv2D(128, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c6)

u7 = tf.keras.layers.Conv2DTranspose(64, (2, 2), strides=(2, 2), padding='same')(c6)
u7 = tf.keras.layers.concatenate([u7, c3])
c7 = tf.keras.layers.Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(u7)
c7 = tf.keras.layers.Dropout(0.2)(c7)
c7 = tf.keras.layers.Conv2D(64, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c7)

u8 = tf.keras.layers.Conv2DTranspose(32, (2, 2), strides=(2, 2), padding='same')(c7)
u8 = tf.keras.layers.concatenate([u8, c2])
c8 = tf.keras.layers.Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(u8)
c8 = tf.keras.layers.Dropout(0.1)(c8)
c8 = tf.keras.layers.Conv2D(32, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c8)

u9 = tf.keras.layers.Conv2DTranspose(16, (2, 2), strides=(2, 2), padding='same')(c8)
u9 = tf.keras.layers.concatenate([u9, c1], axis=3)
c9 = tf.keras.layers.Conv2D(16, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(u9)
c9 = tf.keras.layers.Dropout(0.1)(c9)
c9 = tf.keras.layers.Conv2D(16, (3, 3), activation='relu', kernel_initializer='he_normal', padding='same')(c9)
```

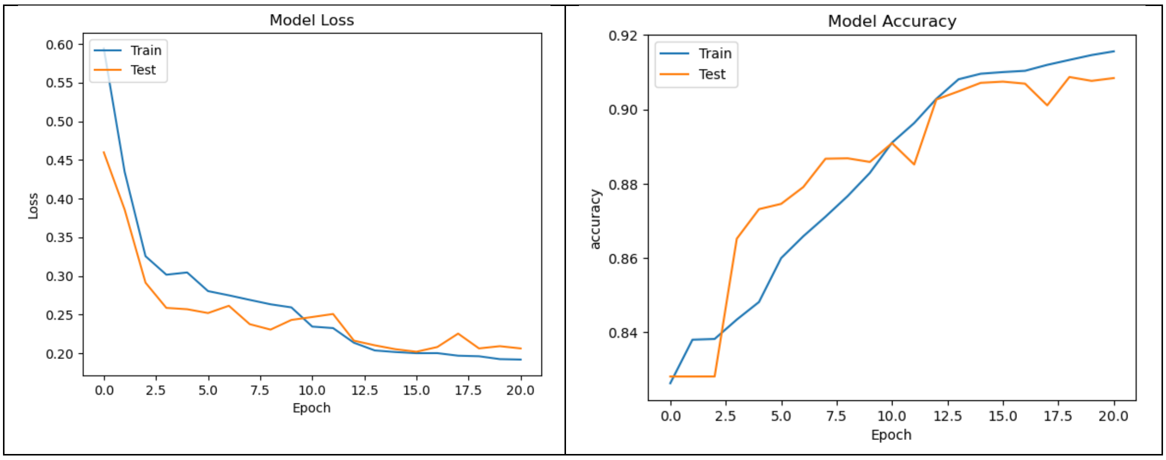
CH 5 Figure 22: Code U-net with python.



CH 5 Table 2: Model 1

4.2 Model 2

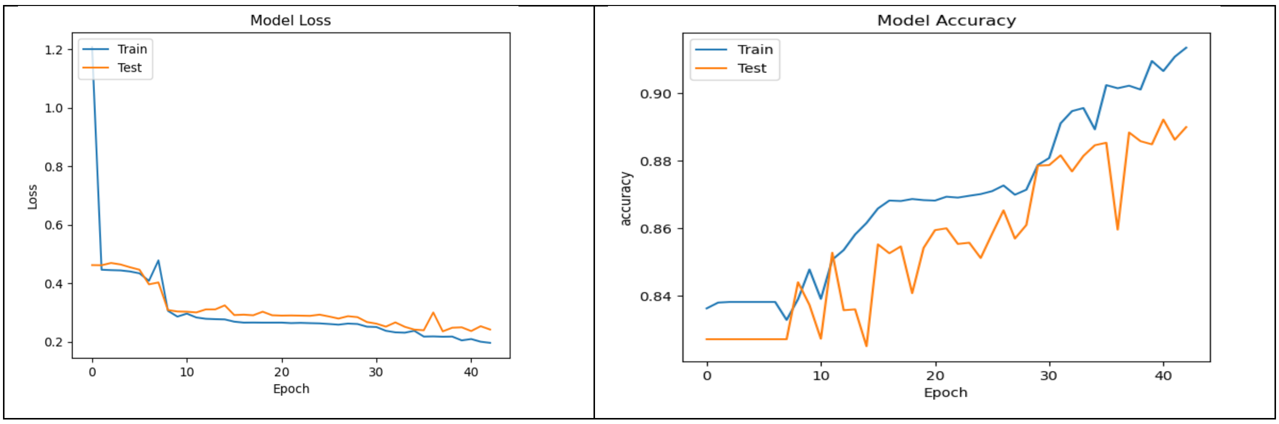
Models were trained using a weighted combination binary_crossentropy. Adam was used for parameter optimization. All the models were trained for 25 epochs total. Start as **32 filter.**



CH 5 Table 3: Model 2

4.3 Model 3

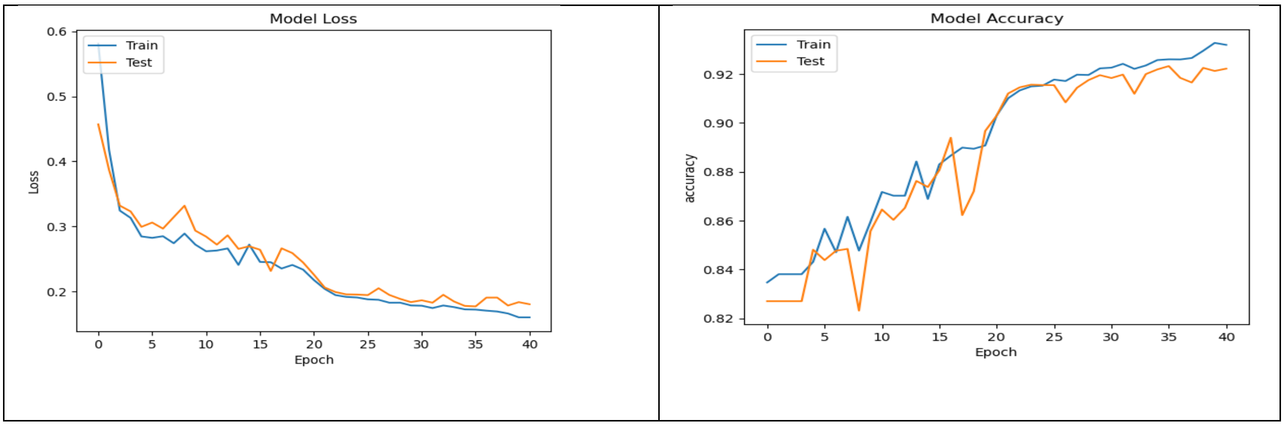
Models were trained using a weighted combination binary_crossentropy. Adam was used for optimization (starting value 0.003). All the models were trained for 45 epochs total. Patch size = 10, start filter as 16.



CH 5 Table 4: Model 3

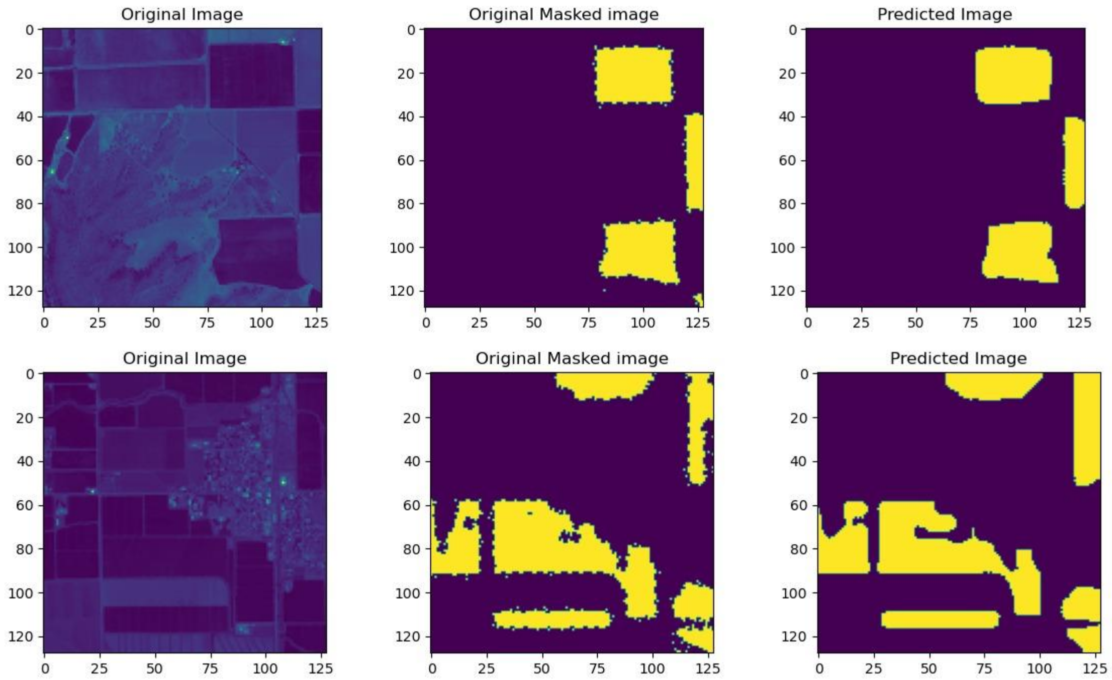
4.4 Model 4

Models were trained using a weighted combination binary_crossentropy. Adam was used for optimization. All the models were trained for 41 epochs total. Patch size = 16, start filter as 16.



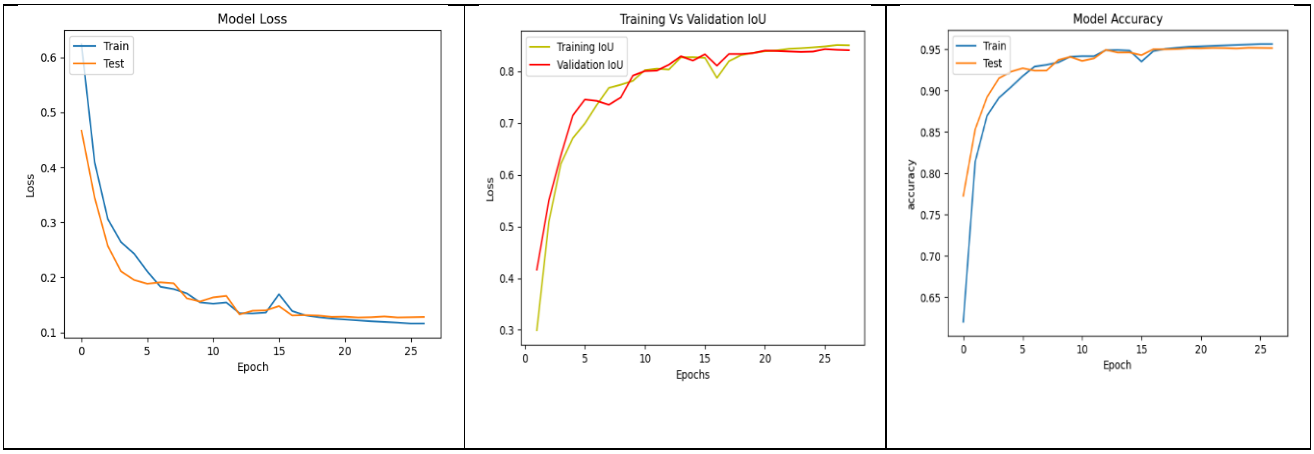
CH 5 Table 5: Model 4

After show the accuracy of model 4. We saw that the model predicts perfectly.



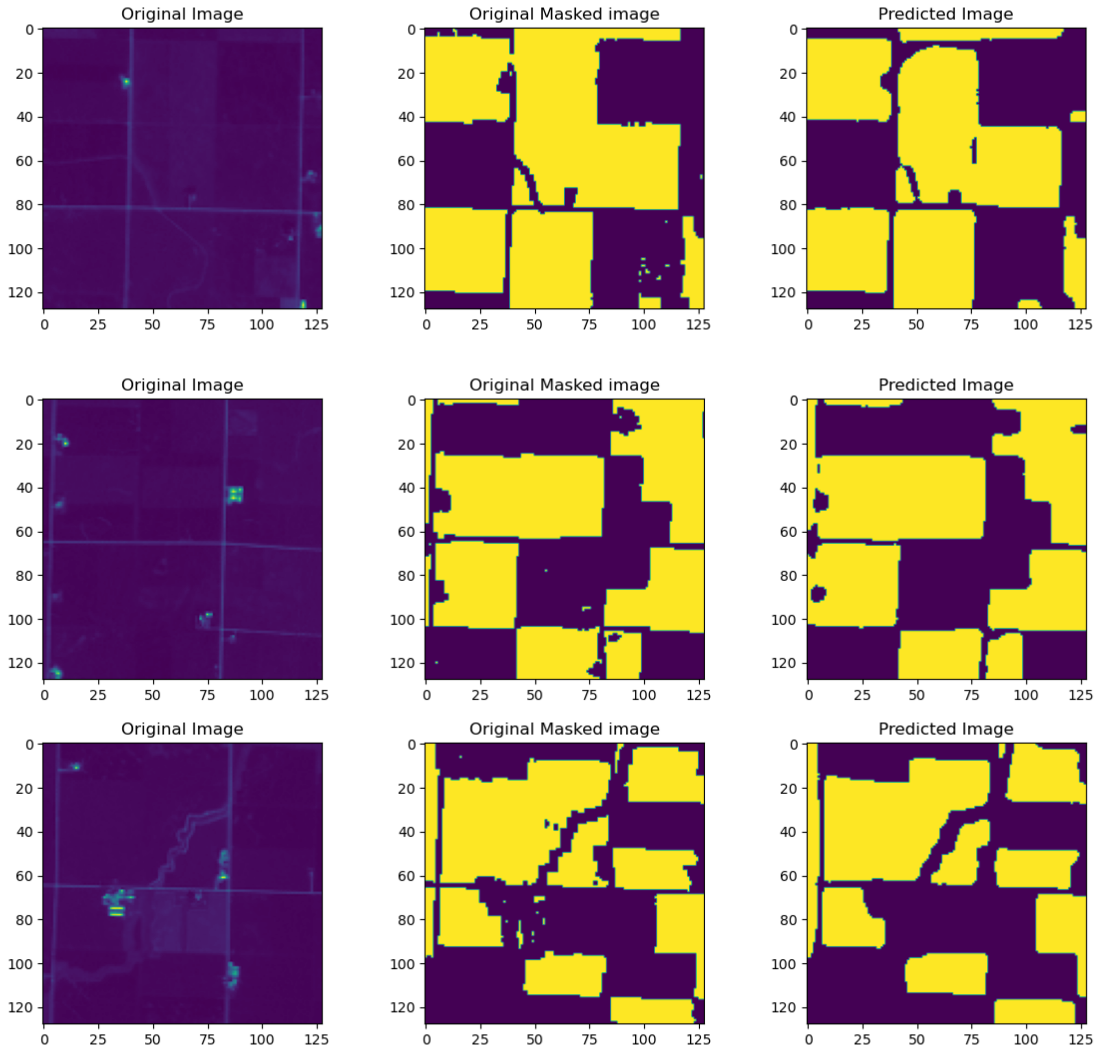
CH 5 Figure 23: Model 4 prediction accuracy

4.5 Experiment 1: This Experiment doing for Corn model with Sentinel-2 only. Models were trained using a weighted combination binary_crossentropy. Adam was used for optimization. All the models were trained for 27 epochs total. Patch size = 16, start filter as 16.



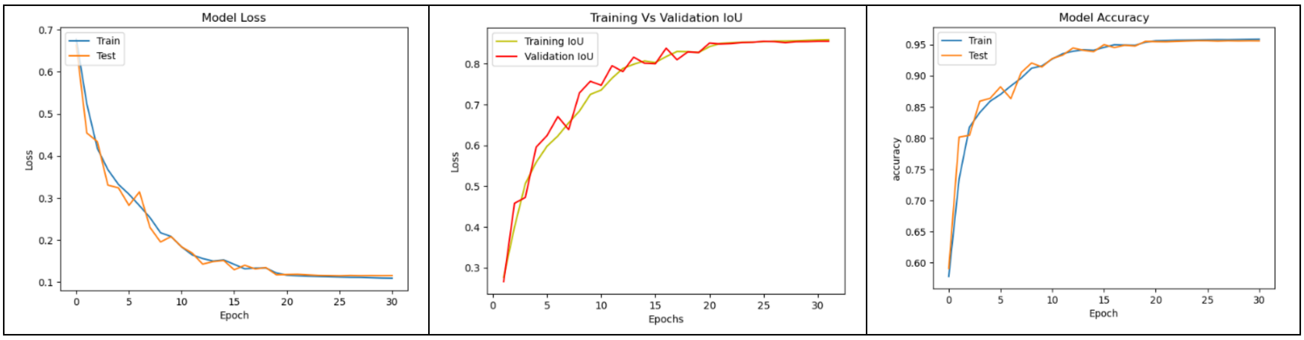
CH 5 Table 6: Experiment 1 for corn model with S2.

As we show this model gives loss: 0.1162 - accuracy: 0.9560 - jaccard_coef: 0.8498. This result is good. Now, we predict it in a new region. We saw that the model predicts perfectly as showed in Table12.



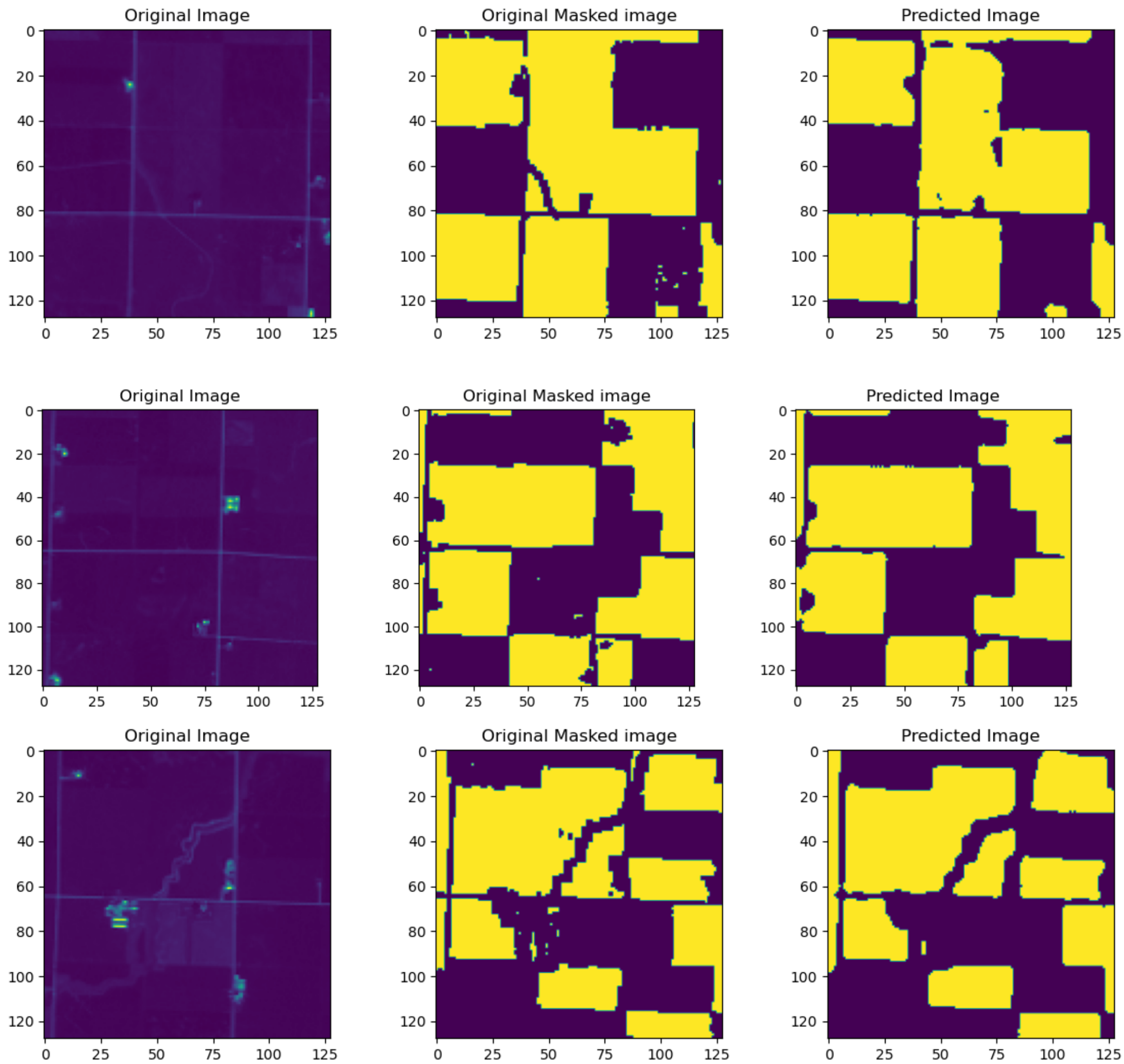
CH 5 Figure 24: Experiment 1 Corn Model Prediction on new region using S2.

4.6 Experiment 2: this Experiment is done for Corn model with Sentinel-2 & Sentinel-1 only. We chose the same hyperparameter of Sentinel-2model and compared them. Models were trained using a weighted combination binary_crossentropy. Adam was used for optimization. All the models were trained for 27 epochs total. Patch size = 16, start filter as 16.



CH 5 Table 7: Experiment 2 for corn model with S1 and S2

As we show this model gives loss: 0.1090 - accuracy: 0.9587 - jaccard_coef: 0.8586. This result is good. Now, we predict it in a new region.



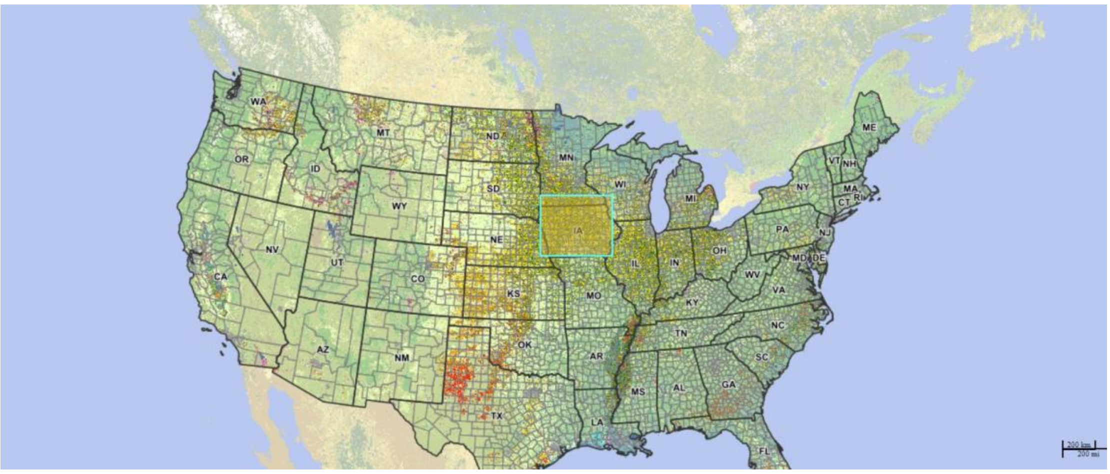
CH 5 Figure 25: Experiment 2 Corn model prediction on new region using S1 and S2.

4.7 Experiment 3: we are done also some experiments. We find that including S1 marginally improves performance. S1 features are indicative of surface scattering, and it may be that crop types within small fields do not have enough differentiating signal in these bands. We find including additional indices such as NDVI, NVWI and NDMI marginally improve performance.

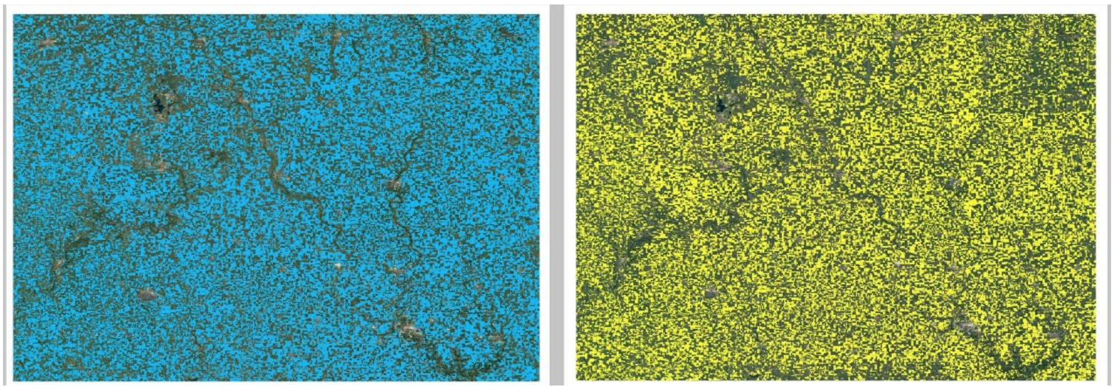
Month	S2	S2&S1	S1&S2&NDVI&NVWI&NDMI
May	0.090%	1.62%	0.39%
June	7.96%	27.12%	30.17%
September	8.91%	27.85%	17.65%
August	84.98%	85.86%	86.02%

CH 5 Table 8: Experiment 3 percentage of Jaccard coefficient.

Experiment in IOWA State.



CH 5 Figure 26: IOWA State



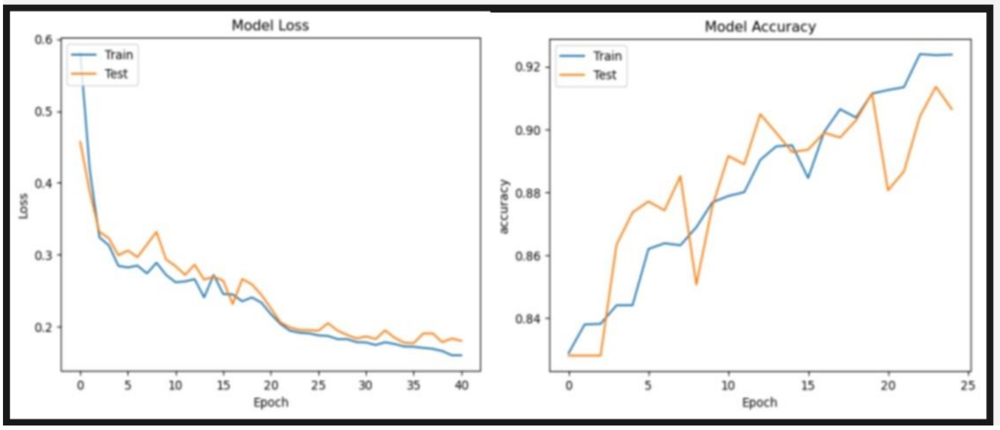
CH 5 Figure 27: ref & our prediction image.

After Experiment in IOWA State, we show good result so we can tell that the model we build it in small area (USA) can predict large area (IOWA State).

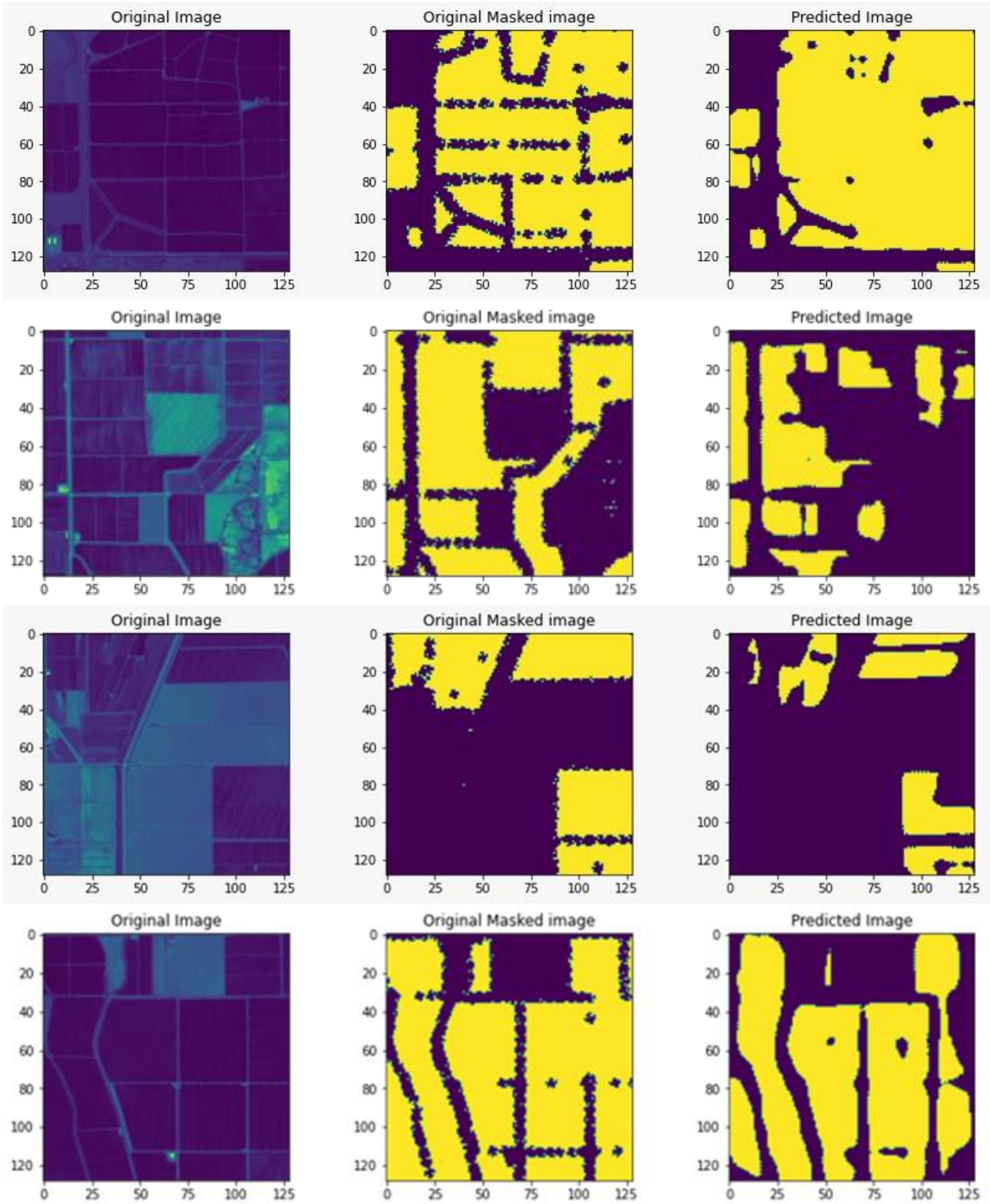
6	Tile ID	Satellite	BBOX Geo Coordinates	Acquision Date	Area [Km2]	Crop Classified [YES/NO]	Drone Photo Available[Yes/No]	Location	Accuracy
7	1	Sentinel-2 L2C	[-95.48749,42.174604,-93.963382,43.209931]	2021-09-16	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile1	51%
8	1	Sentinel-1 RGB	[-96.563986,43.210931,-95.184815,44.22824]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile1	51%
9	2	Sentinel-2 L2C	[-94.547589,42.821105,-93.064029,43.828112]	2021-08-12	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile2	83%
10	2	Sentinel-1 RGB	[-95.48749,43.209931,-94.04777,44.22824]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile2	83%
11	3	Sentinel-2 L2C	[-93.031774,43.352907,-91.622541,44.275457]	2021-9-16	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile3	49.00%
12	3	Sentinel-1 RGB	[-91.832334,43.272223,-90.447679,44.24425]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile3	49.00%
13	4	Sentinel-2 L2C	[-93.031774,43.352907,-91.622541,44.275457]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile4	33%
14	4	Sentinel-1 RGB	[-93.062917,43.292952,-91.603469,44.24425]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile4	33%
15	5	Sentinel-2 L2C	[-93.031774,43.352907,-91.622541,44.275457]	2021-9-18	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile5	8%
16	5	Sentinel-1 RGB	[-91.832334,43.272223,-90.447679,44.24425]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile5	8%
17	6	Sentinel-2 L2C	[-96.797496,42.335963,-95.495598,43.310953]	2021-08-12	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile6	85%
18	6	Sentinel-1 RGB	[-96.65735,42.337993,-95.256275,43.310953]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile6	85%
19	7	Sentinel-2 L2C	[-94.547589,42.821105,-93.064029,43.828112]	2021-08-12	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile7	87%
20	7	Sentinel-1 RGB	[-95.495598,42.335963,-94.099953,43.338927]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile7	87%
21	8	Sentinel-2 L2C	[-95.48749,42.174604,-93.963382,43.209931]	2021-9-23	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile8	30%
22	8	Sentinel-1 RGB	[-94.18824,42.339008,-92.847542,43.338927]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile8	30%
23	9	Sentinel-2 L2C	[-94.245705,41.449976,-92.886334,42.457684]	2021-8-14	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile9	83%
24	9	Sentinel-1 RGB	[-93.015414,42.380614,-91.622541,43.352907]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile9	83%
25	10	Sentinel-2 L2C	[-91.591729,42.597076,-90.448838,43.460411]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile10	22%
26	10	Sentinel-1 RGB	[-91.856384,42.342053,-90.510184,43.350911]	2021-9-28	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile10	22%
27	11	Sentinel-2 L2C	[-96.797496,42.335963,-95.495598,43.310953]	2021-09-26	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile11	19%
28	11	Sentinel-1 RGB	[-96.65735,42.337993,-95.256275,43.310953]	2021-9-16	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile11	19%
29	12	Sentinel-2 L2C	[-94.547589,42.821105,-93.064029,43.828112]	2021-08-12	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile12	53%
30	12	Sentinel-1 RGB	[-95.495598,42.335963,-94.099953,43.338927]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile12	53%
31	13	Sentinel-2 L2C	[-93.031774,43.352907,-91.622541,44.275457]	2021-8-14	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile13	82%
32	13	Sentinel-1 RGB	[-94.245705,41.449976,-92.886334,42.457684]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile13	82%
33	14	Sentinel-2 L2C	[-94.245705,41.449976,-92.886334,42.457684]	2021-8-14	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile14	80%
34	14	Sentinel-1 RGB	[-93.053677,41.487751,-91.655494,42.459703]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile14	80%
35	15	Sentinel-2 L2C	[-91.832334,41.111423,-89.644719,42.762147]	2021-09-10	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile15	56%
36	15	Sentinel-1 RGB	[-91.795936,41.434235,-90.460733,42.443489]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile15	56%
37	16	Sentinel-2 L2C	[-96.65735,41.438342,-95.242526,42.337993]	2021-08-22	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile16	74%
38	16	Sentinel-1 RGB	[-96.640906,40.501011,-95.316692,41.522731]	2021-8-22	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile16	74%
39	17	Sentinel-2 L2C	[-95.601002,40.57539,-94.196449,41.695232]	2021-09-16	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile17	26%
40	17	Sentinel-1 RGB	[-95.388459,40.546941,-94.045013,41.535067]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile17	26%
41	18	Sentinel-2 L2C	[-95.126611,41.269044,-93.510885,42.479239]	2021-08-14	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile18	66%
42	18	Sentinel-1 RGB	[-94.196449,40.580325,-92.90795,41.695232]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile18	66%
43	19	Sentinel-2 L2C	[-94.245705,41.449976,-92.886334,42.457684]	2021-08-14	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile19	72%
44	19	Sentinel-1 RGB	[-93.01269,40.582412,-91.652746,41.535067]	2021-8-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile19	72%
45	20	Sentinel-2 L2C	[-91.832334,41.111423,-89.644719,42.762147]	2021-09-10	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile20	33%
46	20	Sentinel-1 RGB	[-91.795936,41.434235,-90.460733,42.443489]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile20	33%
47	8_new	Sentinel-2 L2C	[-95.48749,42.174604,-93.963382,43.209931]	2021-9-1	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile8_New	68%
48	8_new	Sentinel-1 RGB	[-94.18824,42.339008,-92.847542,43.338927]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile8_New	68%
49	11_new	Sentinel-2 L2C	[-95.48749,42.174604,-93.963382,43.209931]	2021-07-23	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile11_New	82%
50	11_new	Sentinel-1 RGB	[-96.65735,42.337993,-95.256275,43.310953]	2021-7-26	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile11_New	82%
51	17_new	Sentinel-2 L2C	[-95.388459,40.546941,-94.045013,41.535067]	2021-09-06	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile17_New	15%
52	17_new	Sentinel-1 RGB	[-95.388459,40.546941,-94.045013,41.535067]	2021-9-4	12000	YES		https://ai-modelling-lap.notebook.us-east-2.sagemaker.aws/tree/GUI/Dataset/Tile17_New	15%

CH 5 Figure 28: accuracy of model on IOWA state.

4.8 Experiment 4: this Experiment is doing for Rice model Sentinel-2 only.

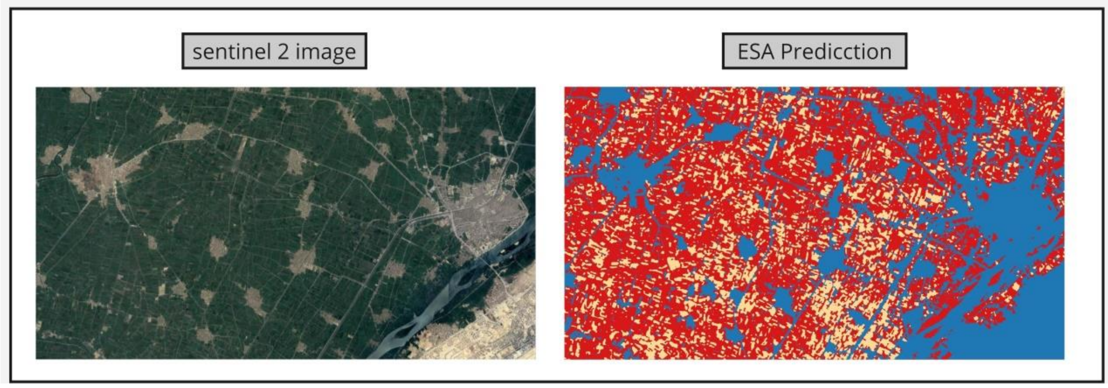


CH 5 Figure 29 Loss & Accuracy for Rice model. ٦

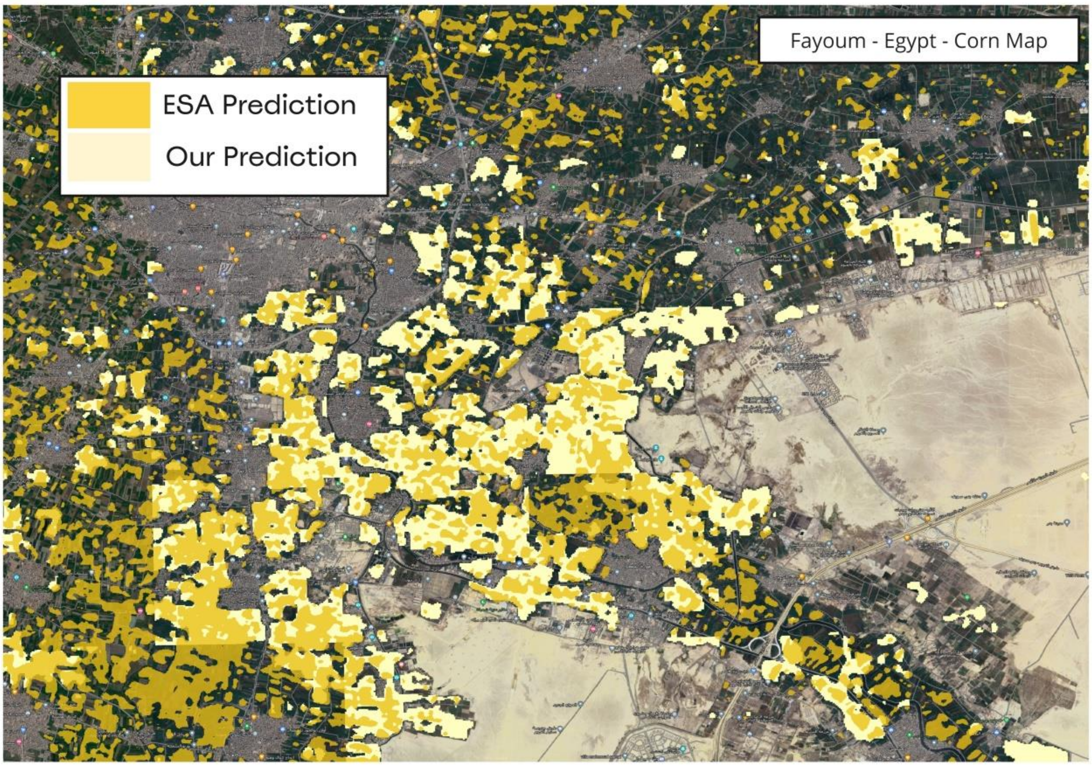


CH 5 Figure 30: Rice Model Prediction Accuracy.

4.9 Experiment 5 In Egypt: this Experiment doing for Corn model Sentinel-2 & Sentinel-1.



CH 5 Figure 31: Egypt-Corn-map.



CH 5 Figure 32: Fayoum-Egypt_Corn Map

As we show, the model predicted Egypt rejoin, but the accuracy is not sufficient because model trained in USA rejoin.

5. Conclusion

In conclusion, with ongoing global population growth and the corresponding rise in demand for food, feed, fiber, and fuel, agriculture plays a key role in sustaining future livelihoods. However, competition for land within the agricultural sector, as well as with other land uses, can lead to negative impacts on ecosystems, such as deforestation and biodiversity loss. The Global Rice map, and other similar tools, is an interactive platform that allows for the worldwide distribution of rice to be showcased. This map plays a vital role in forecasting, monitoring, and assessing rice production at local, regional, and national levels. By providing early access to the spatial distribution of rice, it enables better agricultural management practices, supports policy formulation, and helps to address food security concerns. Rice is a major crop that feeds more than half of the world's population and covers over 12% of the world's cultivated land. Its cultivation has far-reaching effects on water consumption and global warming. Therefore, accurate and timely mapping of rice cultivation is critical for agricultural systems and food security. Remote sensing technology provides an efficient and reliable means of capturing essential agricultural information over large areas. By utilizing this technology, researchers and organizations have made significant strides in improving the accuracy and timeliness of crop type mapping, particularly for rice, almond, and corn crops in California and the USA. The main goal of crop type classification is to create a map or image where each point is classified by what is growing there. Studies have confirmed that remote sensing technology, including different types of vegetation indices, such as the normalized difference vegetation index (NDVI) and the normalized difference water index (NDWI), provide an accurate, fast, and low-consumption technical approach for paddy rice mapping. In conclusion, remote sensing technology has shown great potential in improving agricultural monitoring techniques, particularly in mapping crop types such as rice, almond, and corn. These advancements have facilitated better agricultural management practices, supported policy formulation, and helped address food security concerns, ultimately contributing to the sustainable development of agriculture and its critical role in sustaining future livelihoods. The future plan will be:

- Train Model in July and September.
- Develop our model to predict any rejoin in world.
- Release global Map in 2023.

6 Agriculture patterns detection from Aerial images

6.1. Abstract

Precision agriculture, an emerging field, aims to enhance crop production efficiency using technological tools. Aerial imagery, including satellite and drone images, has become an indispensable asset for various applications, including crisis management. Deep learning methods have significantly advanced relief efforts through recognition, detection, and other non-trivial applications. This research focuses on the application of image semantic segmentation to identify anomalies in crop patterns, such as weeds, dryness, and water locations. By labeling every pixel, this technique enables precise detection of affected areas in the field. Various models and techniques for image segmentation are compared to determine their effectiveness in accurately identifying anomalies. The study also explores the integration of remote sensing, weed management, decision support systems, and traceability to optimize crop management practices. Additionally, the utilization of multispectral images obtained from unmanned aerial systems (UASs) is being investigated. The comprehensive framework developed combines aerial imagery, deep learning, and other agricultural technologies, with the goal of improving crop yield while minimizing resource utilization. The results of this research have far-reaching implications for the agricultural industry, providing a pathway towards sustainable and efficient crop production practices . This study contributes to the advancement of precision agriculture, enabling the detection of anomalies in crop patterns through the synergy of aerial imagery and deep learning techniques. The application of these methods holds exciting prospects for optimizing agricultural practices, ultimately leading to increased productivity and resource efficiency.

6.2. Introduction

Agriculture has always played a vital role in the global economy, sustaining the world's population, and supporting numerous countries' economies. However, the ever-growing demand for food, coupled with environmental challenges and limited resources, has necessitated the development of more efficient and sustainable farming practices. In this context, precision agriculture has emerged as a promising solution to optimize crop production, reduce resource wastage, and mitigate environmental impact. Precision agriculture leverages advanced technologies, including computer vision and deep learning, to enable accurate and detailed monitoring of crops and fields. By harnessing the power of computer vision, farmers and agricultural experts can gain valuable insights into plant health, growth patterns, and field anomalies, facilitating timely intervention and informed decision-making. One of the primary motivations behind the pursuit of precision agriculture is the need to enhance productivity and increase crop yields. With global food demand projected to rise significantly in the coming years, it is imperative to adopt strategies that can boost agricultural output without compromising sustainability. Computer vision techniques provide a non-intrusive and scalable approach to gather data on plant conditions, enabling farmers to optimize resource allocation, tailor treatments, and improve overall crop management practices. Field anomalies pose significant challenges to farmers, as they can negatively impact crop health and productivity. Weeds, for instance, compete with crops for resources and can significantly reduce yields if not promptly identified and controlled. Similarly, issues such as drought stress, nutrient deficiencies, and planter skips can result in irregular plant growth and hinder optimal crop development. Detecting and addressing these anomalies in a timely manner is crucial for minimizing losses and ensuring efficient resource utilization. The traditional approach of manual field scouting and inspection by experts is time-consuming, subjective, and often insufficient for large-scale agricultural operations. Computer vision technology offers a promising alternative, enabling automated and objective analysis of plant images captured from drones, satellites, or ground-based sensors. By leveraging machine learning algorithms, computer vision systems can detect and classify weeds, identify stress patterns, and quantify the extent of planter skips, providing farmers with actionable information for targeted interventions. Furthermore, the consequences of undetected or inadequately addressed field anomalies can extend beyond reduced yields. Ineffective weed control strategies may lead to the development of herbicide-resistant weed populations, jeopardizing long-term crop productivity and increasing reliance on costly interventions. Similarly, overlooking signs of drought stress or nutrient deficiencies can result in irreversible damage to crops, impacting not only the current season's harvest but also the long-term sustainability of agricultural systems. In this research paper, we explore the potential of computer vision techniques in enabling precision agriculture and addressing field anomalies. We aim to provide a comprehensive overview of the advancements in agricultural computer vision, discuss the challenges associated with field anomaly detection, and present potential solutions and research directions. By harnessing the power of computer vision technology, we can revolutionize crop monitoring, optimize resource utilization, and mitigate the risks posed by field anomalies, ultimately contributing to sustainable and efficient agriculture practices on a global scale.

6.3. Literature Review

6.3.1. Introduction

In this section, we will explore the historical progression of semantic segmentation, examining key milestones and breakthroughs that have shaped the field. Additionally, we will delve into the fundamental concepts of artificial intelligence (AI) and machine learning (ML), which underpin semantic segmentation algorithms. Understanding the role of AI and ML in this context is essential for comprehending the advancements made in semantic segmentation. Furthermore, this review will explore the application of deep learning techniques in semantic segmentation. Deep learning, specifically convolutional neural networks (CNNs), has proven highly effective in learning complex visual representations, enabling more accurate and robust semantic segmentation. The relationship between remote sensing and semantic segmentation will also be explored. Remote sensing data, such as satellite imagery and aerial photographs, provide valuable sources of information for semantic segmentation tasks. The integration of remote sensing and semantic segmentation has opened new possibilities for applications in areas such as land cover classification, urban planning, and environmental monitoring. Moreover, an overview of the software tools and frameworks available for semantic segmentation will be provided. These resources offer a range of functionalities, from pre-trained models to evaluation metrics, facilitating the implementation and evaluation of semantic segmentation algorithms. Finally, we will survey existing research in the field, focusing on studies related to our research objectives. By examining the methodologies, datasets, and outcomes of these studies, we can identify gaps and differences that our research aims to address, emphasizing its novelty and contribution to the existing body of knowledge. Through this literature review, we aim to establish a comprehensive understanding of the historical development, core concepts, and current state of semantic segmentation. This foundation will guide our research and contribute to the advancement of semantic segmentation techniques and applications.

6.3.2. Semantic segmentation

Image segmentation is a process of dividing an image into several distinct regions or parts based on certain criteria. However, this process doesn't necessarily involve understanding the image content. On the other hand, semantic segmentation aims to divide an image into meaningful regions based on their semantic or conceptual content. [3]. Semantic segmentation is a computer vision task that involves dividing an image into multiple segments or regions and assigning a class label to each of those segments. The objective of semantic segmentation is to identify every pixel in an image and assign it to a corresponding class label. This means that each segment of the image represents a specific object or part of an object and is labeled accordingly. In other words, the goal of semantic segmentation is to identify the regions of an image that are semantically meaningful, such as objects or parts of objects, and assign each pixel in those regions to a specific class label. This technique is useful in a variety of computer vision applications, such as object detection, image editing, and autonomous navigation. [1] One of the major advances in semantic segmentation has been the use of deep convolutional neural networks (DCNNs). These networks are capable of extracting high-level features from images, which makes them particularly useful in identifying semantically meaningful regions in images. [2]

6.3.3. Early beginnings of semantic segmentation.

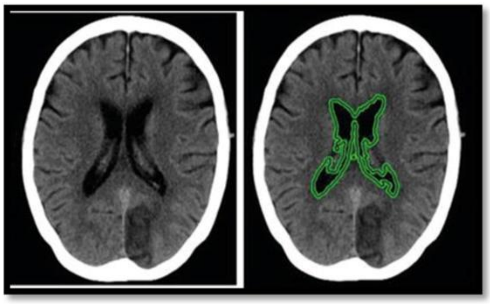
The early beginnings of semantic segmentation can be traced back to the development of computer vision and image processing techniques. [3] Where Segmentation is the process of separation of required information from a data for further processing. [4] In the early stages, the focus was primarily on low-level image analysis tasks such as edge detection and feature extraction. [3]

6.3.3.1. <u>**Active Contour**</u>

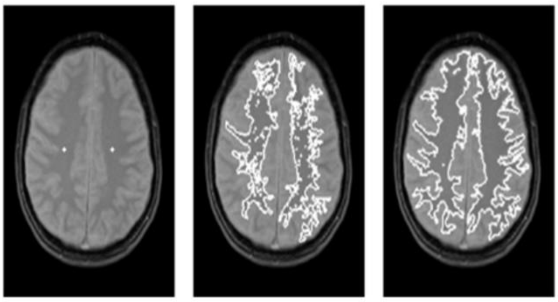
The Active Contour Method is a technique that uses a parametric active contour model based on feature images to extract object contours. Active contours, also known as snakes, are deformable models that use constraints and forces within an image for segmentation. By leveraging these active contours, the method accurately delineates the boundaries of objects of interest. This approach is particularly useful for objects with irregular shapes or complex boundaries, enabling precise segmentation and further analysis. [3] [4]

6.3.3.2. <u>***Region Growing***</u>

The Region Growing Method is a technique that involves selecting seed points and growing regions from these points to adjacent points based on a region membership criterion. This approach exploits the fact that pixels close to each other tend to have similar gray values. It starts with a single pixel as a seed and gradually adds new pixels to the region. In contrast to histogram-based region detection, which lacks spatial information and only provides the distribution of gray levels, region-growing approaches consider the spatial proximity of pixels to determine their membership in a region. By iteratively expanding the region from the seed point, this method can effectively identify connected regions with similar characteristics. [5][6]



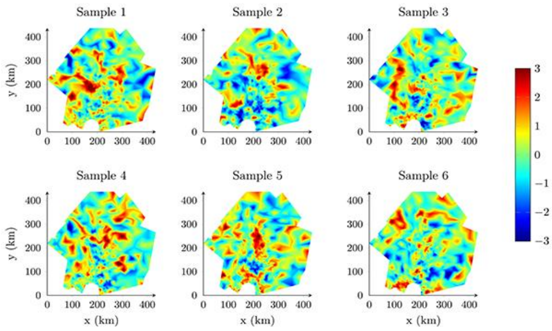
CH 6 Figure1 :Segmentation of brain CT image using active contours. [4]



CH 6 Figure 2 Region (seed) Growing Segmentation. [6]

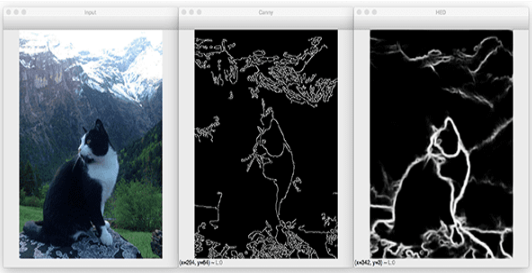
6.3.3.3. <u>Markov Random Field</u>

A Markov Random Field (MRF) is a graphical model where nodes represent random variables and edges capture local influences between them. The connectivity of the graph enables the propagation of these influences globally. MRFs are powerful tools for modeling dependencies and interactions in a probabilistic framework. They find applications in various fields, such as computer vision and machine learning, to capture complex relationships and solve inference problems efficiently. By representing the relationships between variables, MRFs provide a comprehensive understanding of the system being modeled. **[7][8]**



CH 6 Figure 3 :Implementation of a Gaussian Markov random field sampler for forward uncertainty quantification in the Icesheet and Sea-level System Model v4.19 . [8]

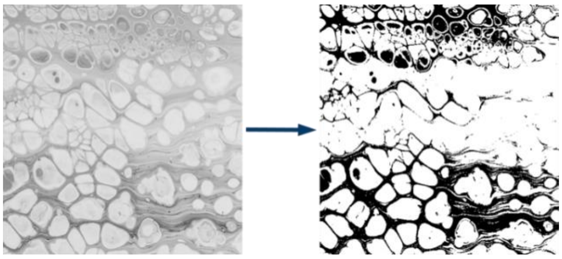
6.3.3.4. <u>Edge Detection</u>



CH 6 Figure 4 :Edge detection via the HED approach with OpenCV and deep learnin.[10]

Edge detection is a fundamental technique in image processing that aims to identify boundaries of objects within an image. It operates by detecting abrupt changes in brightness, which are referred to as discontinuities. These abrupt changes typically indicate the presence of edges or boundaries between different regions in the image. By locating and highlighting these edges, edge detection algorithms facilitate subsequent image analysis tasks such as object recognition, segmentation, and feature extraction. The identification of edges plays a crucial role in various applications, including computer vision, robotics, and medical imaging [9] [10]

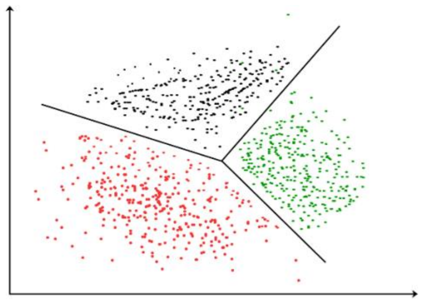
6.3.3.5. Thresholding



CH 6 Figure 5 :Image Thresholding Based on Otsu's Method using OpenCV.1[12]

Segmentation is the process of dividing an image into regions or identifying the contours corresponding to objects. Common properties are used to separate these regions or identify the differences between them. One simple property that pixels in a region can share is their intensity. This leads to thresholding, a segmentation technique where a grayscale image is transformed into a binary image with only two values, 0 and 1, based on a threshold value. Pixels with intensity values above the threshold are assigned the value 1 (white), while the rest are assigned the value 0 (black). Thresholding allows for the separation of light and dark regions, providing a basic but effective method for image segmentation [11] [12].

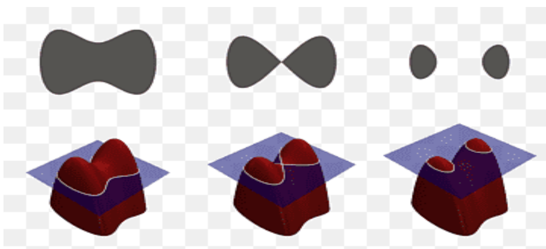
6.3.3.6. Clustering Method



CH 6 Figure 6: Clustering in Machine Learning.- [14]

Clustering is an unsupervised learning method in machine learning that involves drawing inferences from unlabeled datasets. It is an exploratory data analysis technique used to analyze multivariate datasets. The goal of clustering is to divide the dataset into a specific number of clusters, where data points within each cluster share similar characteristics. Clusters are formed by grouping data points together, minimizing the distance between them. The purpose of clustering is to identify and segregate groups with similar traits or patterns. This technique is useful for understanding the inherent structure and relationships within the data, enabling further analysis and insights [13] [14].

6.3.3.7. Level Set Method

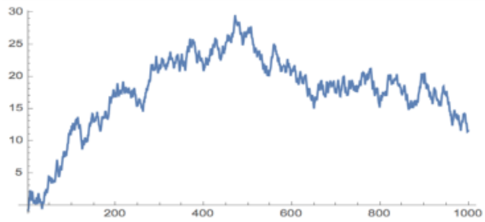


CH 6 Figure 7: Level-set method Level set Shape Volume of fluid method, shape, angle, shape, and topology. [16]

The level-set method is a technique that relies on a continuous level-set function, where the interface or boundary is defined by the zero contour of the function. The level-set function, denoted as ϕ, is a scalar field that represents the distance to the nearest interface or boundary. By using this function, the level-set method provides a powerful framework for tracking and evolving interfaces or boundaries in various applications, including image processing and shape optimization. It allows for efficient and accurate representation and manipulation of complex geometries, making it particularly useful in problems involving shape analysis, object tracking, and image segmentation. [15]

6.3.3.8. Random Walk Method

Random walk theory posits that changes in asset prices, such as stock prices, follow a random pattern. This implies that past prices cannot be reliably used to predict future prices, as stock prices move in an unpredictable manner. Furthermore, random walk theory suggests that the stock market is efficient, meaning that it incorporates and reflects all available information.



CH 6 Figure 8: Plot a Random Walk. [17]

The concept of random walk challenges the notion that traders can time the market or employ technical analysis to identify and profit from patterns or trends in stock prices. Critics of random walk theory argue that stock prices can be predicted using various methods, such as technical analysis. These traders and analysts believe that there are exploitable patterns and trends in stock prices that can be identified and utilized for profitable trading strategies [17].

6.3.4. AI, ML, and Deep Learning: Revolutionizing Image Segmentation

Artificial Intelligence (AI) and Machine Learning (ML) techniques, particularly Deep Learning, have revolutionized image segmentation tasks. Image segmentation involves partitioning an image into distinct regions or objects. AI and ML algorithms can learn from large amounts of labeled data to automatically identify and delineate these regions accurately. Deep Learning, with its deep neural networks, has shown remarkable performance in image segmentation by leveraging its ability to extract complex features and learn hierarchical representations from data. This enables precise and fine-grained segmentation of objects in images.

6.3.4.1. Artificial Intelligence (AI)

Artificial Intelligence (AI) has emerged as a transformative technology with the potential to revolutionize various industries and domains.[18] AI refers to the development of intelligent systems that can perceive reason, learn, and make decisions similar to human intelligence. In recent years, AI has gained significant attention and has become a driving force behind numerous advancements.

6.3.4.2. Machine Learning (ML)

Machine Learning (ML) is a subset of Artificial Intelligence (AI) that focuses on the development of algorithms and models capable of learning from data and making predictions or decisions without explicit programming. ML techniques have revolutionized various industries and domains by enabling computers to automatically learn patterns, extract insights, and perform complex tasks. In the field of image segmentation, ML techniques have played a pivotal role in improving accuracy and efficiency. ML algorithms learn from labeled datasets to identify and classify different regions or objects within images. They leverage statistical techniques, pattern recognition, and optimization algorithms to analyze pixel-level information and make accurate segmentation decisions.ML-based image segmentation methods include clustering algorithms, decision trees, support vector machines (SVM), and random forests, among others. These algorithms can analyze feature spaces, learn from training data, and make predictions on new, unseen images. ML techniques allow for the automated and precise delineation of objects, eliminating the need for manual intervention. The integration of ML techniques in image segmentation has revolutionized the field, enabling automated and accurate solutions across various industries. With ongoing advancements and the availability of large-scale labeled datasets, ML-based image segmentation holds great potential for further improvement and broader applications in the future.

6.3.4.3. deep learning (DL)

Deep Learning (DL) has emerged as a powerful subset of Machine Learning (ML) that utilizes artificial neural networks with multiple layers to extract hierarchical representations and learn complex patterns from data. DL has revolutionized various fields, including image segmentation, by significantly improving accuracy and performance.

6.3.4.3.1. Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANNs) are a family of machine learning models that are inspired by the biological structure of the human brain. These networks consist of a set of units or nodes, which are connected in a way that mimics the neural architecture of the brain. These units, also known as artificial neurons, are aggregated into layers that are connected in a sequential manner. Typically, not all neurons are directly connected, and signals travel from the input layer to the output layer through intermediate layers called hidden layers. The strength of the connections between pairs of neurons is represented as model parameters that can be learned through a process called model training, which is based on a technique called back-propagation. This technique adjusts the individual connection values, also known as weights, to minimize a function that represents how big the errors are that the network commits while performing the task. [3]

6.3.4.3.2. Deep Artificial Neural Networks (DNNs)

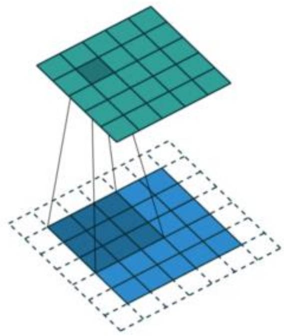
Deep Artificial Neural Networks are a specific type of ANNs that contain more than three layers, including one input layer, one hidden layer, and one output layer. The intuition behind using deep networks is that each layer adds its own level of non-linearity, achieving a higher abstraction capacity that cannot be reached by a single hidden layer. Although theoretically, a shallow network with enough neurons in the hidden layer can represent any function, deep networks work much better in practice. In a DNN, each layer's inputs are linearly combined, and hence cannot produce the non-linearity that can be seen through multiple layers. [19]

6.3.4.3.2. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks are a subfamily of ANNs that are designed to deal with data such as images or audio. CNNs take inspiration from the visual cortex of the brain, where different neurons or groups of neurons are sensitive to different patterns and fire accordingly. Each layer in a CNN is characterized by several parameters. [19]

i. **Kernel size:** the kernel size defines the field of view of the convolution as it refers to the spatial dimensions of the filter (the size is generally the same is all spatial dimensions, but it is not necessary).
ii. **Input depth:** refers to the number of stacked feature maps received from the previous layer. Thus, the filter dimensions in a given convolutional layer will be the spatial dimensions and the input depth.
iii. **Output depth or number of filters:** the size of the filter bank contained in a layer. Each filter will end up generating a feature map. These filter maps will be stacked and passed forward to the next layer. Consequently, the number of filters in one layer is equivalent to the input depth in the next one.
iv. **Stride:** stride is the number of pixels slid along each dimension to calculate consecutive filter outputs.
v. **Padding:** this parameter defines how the border of a sample is handled. Unpadded convolutions will crop away some of the borders if the kernel size is larger than 1.

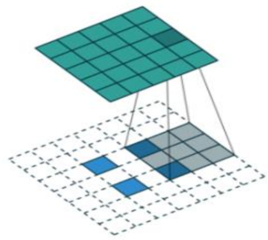
All together, these parameters define a convolution operation.



CH 6 Figure 9: a 2D convolution with a kernel size of 3stride of 1 and padding. [٣]

6.3.4.3.2.1. Transposed convolutional layers

In the field of deep learning, transposed convolution, also known as deconvolution or fractionally strided convolution, is an interesting operation that reverses the effect of a normal convolution. While a regular convolution reduces the spatial dimensions of the input, transposed convolution increases the spatial dimensions of the output while preserving the connectivity pattern of a standard convolution. This is achieved through the use of clever padding techniques on the input data, which enable the expansion of the output spatial dimensions. (Figure 6-10) provides an illustrative example of transposed convolution, demonstrating the increase in spatial dimensions from a 4-dimensional space to a 16-dimensional space. Transposed convolution plays a crucial role in tasks such as image upsampling, image synthesis, and generating high-resolution feature maps in deep learning architectures. [19]

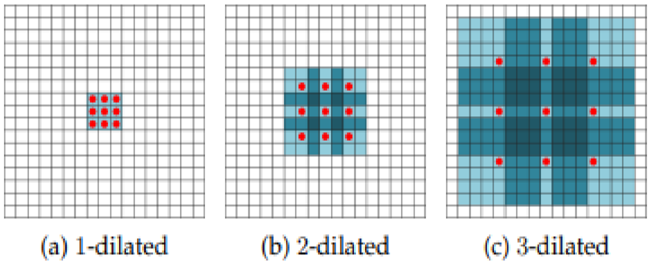


CH 6 Figure 10: Example of 2D transposed convolution. [٣]

Dilated Convolutions

Dilated convolutions, also known as "a-trous convolutions," offer a way to expand the receptive fields of convolutional filters without sacrificing resolution. They are a generalization of Kronecker-factored convolutional filters and utilize unsampled filters. The dilation rate, denoted as "l," controls the upsampling factor. By stacking l-dilated convolutions, the receptive fields grow exponentially while keeping the number of filter parameters linear. This property enables efficient dense feature extraction at any resolution. It's worth noting that typical convolutions can be considered as 1-dilated convolutions. Dilated convolutions are a valuable tool in deep learning architectures for tasks such as image segmentation, where capturing context across different scales is crucial. [20]

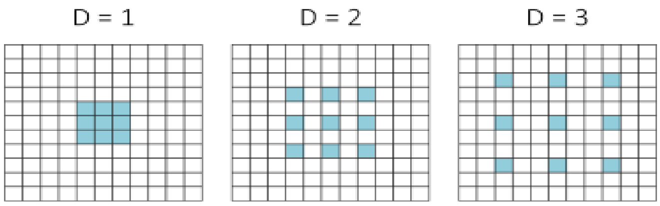
Fig. ٦-11: As shown, dilated convolution filters with various dilation rates: (a) 1-dilated convolutions in which each unit has 3×3 receptive fields, (b) 2-dilated ones with 7 × 7 receptive fields, and (c) 3-dilated convolutions with 15 × 15 receptive fields.



CH 6 Figure 11: Dilated Convolutions [20]

In practical terms, implementing dilated convolutions involves dilating or expanding the filter before performing the convolution operation. This expansion is done by increasing the size of the filter according to the specified dilation rate, while filling the additional elements with zeros. As a result, the filter weights are matched with distant elements that are not adjacent if the dilation rate is greater than one. This allows the network to capture a broader context and incorporate information from a larger receptive field. Figure 6-10 provides visual examples of dilated filters.

Fig. 6-12: Filter elements (green) matched to input elements when using 3×3 dilated convolutions with various dilation rates. From left to right: 1, 2, and 3. [20]

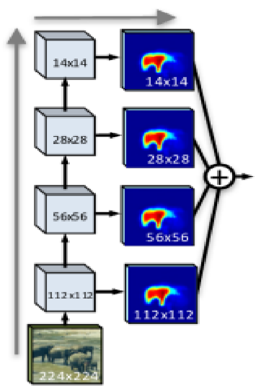


CH 6 Figure 12: examples of dilated filters [20]

6.3.4.3.3. Feature Fusion

To enhance the performance of fully convolutional architectures in image segmentation, feature fusion is employed as a method to incorporate contextual information. This technique involves merging a global feature, obtained from a previous layer, with a more localized feature map extracted from a subsequent layer. Skip connections, seen in architectures like the original FCN, facilitate a late fusion approach by combining feature maps from multiple layers.

Fig. ٦-13: Skip-connection-like architecture, which performs late fusion of feature maps as if making independent predictions for each layer and merging the results. Figure extracted from. [20]



CH 6 Figure 13: Skip-connection-like architecture [20]

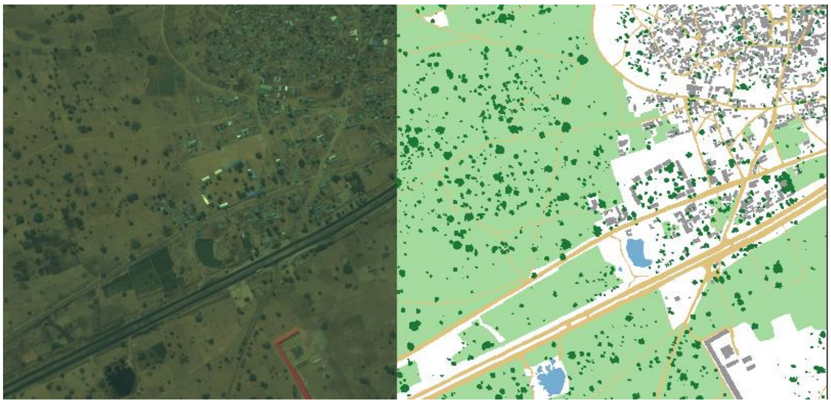
In conclusion, Deep Learning has revolutionized image segmentation by leveraging the power of Artificial Neural Networks, specifically Convolutional Neural Networks. These networks have shown exceptional capability in extracting hierarchical representations and learning complex patterns from image data. Techniques such as transposed convolutional layers, dilated convolutions, and feature fusion have further enhanced the accuracy and efficiency of segmentation models. With continuous research and advancements in Deep Learning, we can expect further improvements in segmentation algorithms, leading to more accurate and robust results.

6.3.5. Remote sensing and UAVs

Remote sensing, which involves acquiring data from a remote sensing device and extracting valuable information from it, plays a crucial role in various applications.[21] Unmanned aerial vehicles (UAVs) equipped with sensors have emerged as a versatile tool in remote sensing, finding applications in agriculture, forestry, mining, and more. While UAVs aim to be a general remote sensing tool, there is a need for tailored data processing and analysis methods specific to different applications. UAV data offers advantages such as high spatial resolution and flexibility in integration with sensors, making them comparable to traditional airborne and spaceborne remote sensing platforms. [22][23] Although satellite imaging and ground-based technology have been studied for crop sensing, they suffer from limitations. Satellite imagery often lacks the spatial resolution required to capture important crop metrics like the leaf area index (LAI), essential for estimating crop growth. Additionally, satellite sensors cannot sense visible light during cloudy conditions, posing a significant challenge. On the other hand, ground-based sensing encounters difficulties when dealing with tall crops like maize, as accessing the fields becomes impractical. [22][23]

6.3.6. Integration of Remote Sensing and Semantic Segmentation

The integration of remote sensing and semantic segmentation has revolutionized our ability to extract valuable information from vast amounts of remote sensing data. Remote sensing, through sources like satellite or UAVs imagery, provides us with a comprehensive view of the Earth's surface. This integration has the potential to transform the way we study and monitor our environment, enabling us to make informed decisions, mitigate risks, and manage resources more effectively.



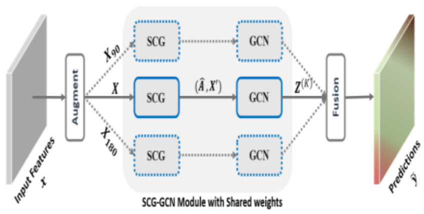
CH 6 Figure 14: illustrates the effectiveness of semantic segmentation in accurately classifying five distinct classes: **b**uildings, roads, trees, crops, and water [24].

6.3.7. Agriculture-Vision Challenge: Methodologies and Results

In the following section, we will present an overview of the methodologies employed by various teams in the Agriculture-Vision challenge. The challenge focused on the development of advanced computer vision models for accurate classification and segmentation of agricultural areas using aerial images. Additionally, we will discuss the dataset provided by the Agriculture-Vision challenge, which will be further detailed in Section 6.2.1 titled "Methodologies: Dataset" The methodologies employed by the participating teams showcased a range of innovative techniques and approaches. These methodologies encompassed various aspects, such as addressing imbalanced class distributions, handling feature divergence between different image channels, exploring fusion techniques for multi-spectral images, incorporating attention and feature extraction blocks, and tackling challenges related to varying shapes and sizes of anomaly patterns.

i. Multi-view Self-Constructing Graph Convolutional Networks with Adaptive Class Weighting Loss

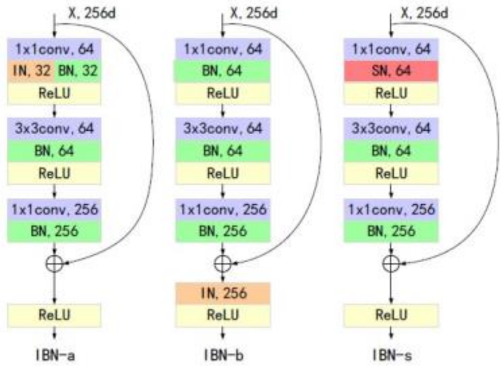
- Introduced a novel method for adaptive class weighting in multi-class segmentation tasks.
- Implemented iterative batch-wise class rectification to address the issue of imbalanced class distribution in training data.
- Computed pixel-frequency for each class to derive adaptive class weights.
- Developed an adaptive multi-class weighting (ACW) loss function to balance positive and negative samples.
- Utilized pertained backbone models, applied data augmentation techniques, and employed specific optimization strategies.



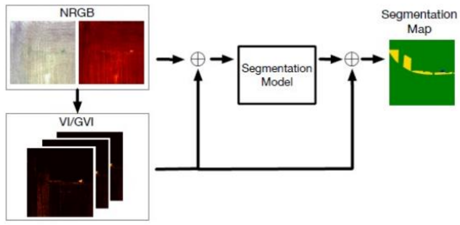
CH 6 Figure 15: Adaptive Class Weighting Loss [25]

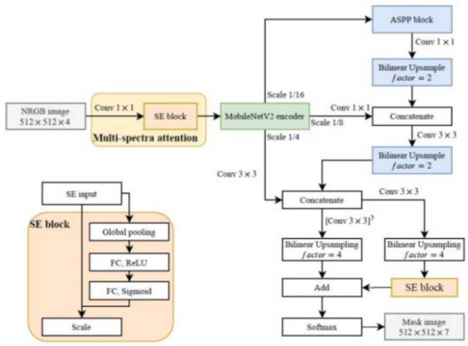
ii. Reducing the feature divergence of RGB and near-infrared images using Switchable Normalization.

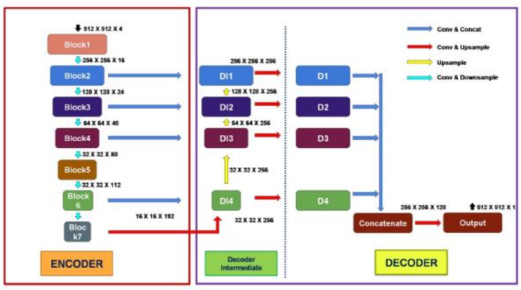
- Focused on addressing the feature divergence between RGB and near-infrared (NIR) images in agricultural aerial image processing.
- Proposed the use of a Switchable Normalization block integrated into the DeepLabV3+ segmentation model.
- Applied the symmetric Kullback-Leibler (KL) divergence measure to significantly reduce the divergence between RGB and NIR channels.
- Introduced a hybrid loss function to handle imbalanced data and optimize the evaluation metric.



CH 6 Figure 16: Figure introduce the IBN-s Block [25]

<table>
<tr>
<td>iii. Team AGR (University of Illinois at Urbana Champaign)
• Applied the focal loss and Lovsz-Softmax functions to handle imbalanced datasets.
• Explored the use of additional input channels, such as the Normalized Difference Vegetation Index (NDVI) and image mask.
• Utilized the ESP Net V2 model as the base model, trained with the Adam optimizer, and introduced dropout layers.</td>
<td></td>
</tr>
<tr>
<td>iv. Team Haossr (Stanford University, Chegg, Inc.)
• Explored effective fusion techniques for multi-spectral agricultural images.
• Proposed a learnable generalized vegetation index module that could be concatenated with the original color channels.
• Introduced an additive group normalization module to ensure smooth training of the model with the generalized vegetation index output.</td>
<td>

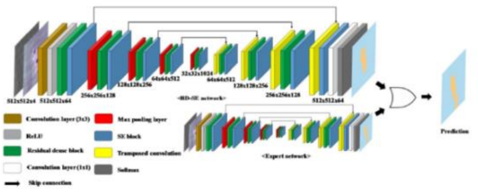
CH 6 Figure 17: illustration of the fusion module for the generalized vegetation index. [25]</td>
</tr>
<tr>
<td>v. Team CNUPR TH2L (Chonnam National University, Chosun University)
• Deployed a Deep Convolutional Encoder-Decoder architecture based on MobileNetV2.
• Incorporated attention blocks to assign the contribution of each spectral channel.
• Utilized ASPP blocks for multi-scale feature extraction and squeeze-excitation blocks for upsampling the feature map to the original input size.</td>
<td>

CH 6 Figure 18: Team CNUPR TH2L: pipeline [25]</td>
</tr>
<tr>
<td>vi. Team TeamTiger (SGGS Institute of Engineering and Technology) [25]
• Developed an encoder-decoder architecture using EfficientNet as the backbone and a feature pyramid decoder.
• Addressed challenges related to varying shapes and sizes of anomaly patterns, imbalanced class distribution, and overlapping labels.
• Employed the Jaccard loss as the objective function for training the model and utilized the Adam optimizer.</td>
<td>

CH 6 Figure 19: Team TeamTiger: proposed encoder-decoder architecture</td>
</tr>
</table>

vii. Team DSSC

- Adopted the Residual DenseNet with Squeeze-and-Excitation blocks (RD-SE) as the base model for semantic segmentation.
- Implemented residual dense blocks and skip connections to compensate for spatial loss during feature extraction.
- Utilized Squeeze-and-Excitation blocks to recalibrate channel-wise feature responses.



CH 6 Figure 20: Team DSSC: Residual DenseNet with Expert Network architecture. [25]

In the Agriculture-Vision challenge, several teams showcased their innovative methodologies to tackle the complexities of the task. Each team's approach was evaluated based on the modified mean Intersection over Union (mIOU) matrix, which will be discussed in the methodology section. The following section provides a summary of the results achieved by each team and highlights the effectiveness of their respective methodologies. These results serve as a valuable benchmark for comparing and evaluating the performance of my own method, which will be presented in the results section. By comparing my results with those of other teams, a comprehensive assessment of the effectiveness and competitiveness of my approach can be made.

Submission	modified mIoU	Back-ground	Cloud shadow	Double plant	Planter skip	Standing water	Water-way	Weed cluster
DSSC	63.9	80.6	56	57.9	57.5	75	63.7	56.9
seungjae	62.2	79.3	44.4	60.4	65.9	76.9	55.4	53.2
yjl9122	61.5	80.1	53.7	46.1	48.6	76.8	71.5	53.6
SCG_Vision	60.8	80.5	51	58.6	49.8	72	59.8	53.8
AGR	60.5	80.2	43.8	57.5	51.6	75.3	66.2	49.2
SYDu	59.5	81.3	41.6	50.3	43.4	73.2	71.7	55.2
agri	59.2	78.2	55.8	42.9	42	77.5	64.7	53.2
TJU	57.4	79.9	36.6	54.8	41.4	69.8	66.9	52
celery030	55.4	79.1	38.9	43.3	41.2	73	61.5	50.5
stevenwudi	55	77.4	42	54.4	20.1	69.5	67.7	53.8
PAII	55	79.9	38.6	47.6	26.2	74.6	62.1	55.7
agrichallenge12	54.6	80.9	50.9	39.3	29.2	73.4	57.8	50.5
hui	54	80.2	41.6	46.4	20.8	72.8	64.8	51.4
shenchen616	53.7	79.4	36.7	56.3	21.6	67	61.8	52.8
NTU	53.6	79.8	41.4	49.4	13.5	73.3	61.8	56
tpys	53	81.1	50.5	37.1	25.9	67.4	58.7	50.1
Simple	52.7	80.2	40	45.2	24.6	70.9	57.6	50.4
Ursus	52.3	78.9	36.3	37.8	34.4	69.3	57.1	52.3
liepieshov	52.1	77.2	40.2	46	16	71.3	62.9	51.1
Lunhao	49.4	79.5	40.4	38.8	10.5	69.4	58.3	49.1
tetelias-mipt	49.2	80.4	37.8	34.8	4.6	70.6	62.5	53.8
Dataloader	48.9	79.1	42	35.8	9.1	68.7	56.7	51.3
Hakjin	46.4	78.6	32	38.3	1.8	66.2	58	49.9
JianyuTANG	44.6	78.1	37.9	31.8	15.4	47.3	54.8	46.9
Haossr	43.9	79.2	21.4	28.1	2.7	67.5	56.4	52.3
rpartsey	41.5	72.5	21.6	36.2	9.1	59.7	40.7	50.6
TeamTiger	40.8	75.2	26.1	40.1	9.9	48	37.1	49.5
Chaturlal	40.7	77.7	23	20.4	5	55	51	52.9

CH 6 Table 1: represents the results of different teams based on modified mIOU matrix [25]

6.4. Methodology

In this section, we provide a comprehensive overview of the methods utilized for semantic segmentation, describe the dataset employed, and detail the evaluation metrics used to assess the accuracy of our models. Our study harnesses the power of deep learning techniques, specifically by employing Fully Convolutional Networks (FCNs), to capture intricate image features and generate spatially coherent segmentations. The dataset utilized in our research is carefully curated, comprising diverse images that have been meticulously annotated at the pixel level. To ensure the quality and suitability of the dataset, rigorous preprocessing techniques are applied, including the division of data into distinct training, validation, and testing sets. For evaluating the performance of our models, we employ well-established evaluation metrics such as Pixel Accuracy, Intersection over Union (IoU), and F1 score. These metrics provide a comprehensive assessment of the segmentation results, considering both pixel-level accuracy and the spatial overlap between predicted and ground truth masks. We aim to achieve accurate and reliable semantic segmentation results.

6.4.1. Dataset

The agriculture-Vision dataset, provided by the Agriculture Vision Organization, offers a valuable resource for researchers and practitioners in the field of computer vision for agriculture. With its focus on aerial farmland images and semantic segmentation of agricultural patterns, this dataset presents unique characteristics and challenges that set it apart from other image datasets. One notable aspect of the agriculture-Vision dataset is the unprecedented aerial image resolution it offers, with pixel resolutions as high as 10 cm per pixel. This level of detail enables researchers to analyze farmland images at a granular level and extract valuable insights related to crop health, field conditions, and agricultural practices. In addition to high-resolution RGB channels, the dataset also includes Near-infrared (NIR) channels, providing an expanded range of image data for analysis. This multi-channel information allows for more comprehensive and accurate assessment of various agricultural phenomena, such as plant health, moisture content and nutrient levels. [25]

A key highlight of the Agriculture-Vision dataset is the presence of challenging annotations of multiple agricultural anomaly patterns. These annotations have been meticulously created by professional agronomists with domain expertise, ensuring a high level of accuracy and reliability. The dataset encompasses eight distinct types of field anomaly patterns that have significant impacts on crop conditions and final yield. These patterns include. [25]

Double Plant: Identifying instances where two plants grow closely together enables optimization of plant spacing, avoiding resource competition, and maximizing crop yield.

Dry Down: Detecting areas with reduced moisture content helps optimize irrigation practices, ensuring proper water availability for crops and mitigating the risk of stress-induced yield losses.

Nutrient Deficiency: Recognizing areas displaying nutrient deficiency symptoms allows for targeted fertilization, minimizing costs and environmental impact while promoting optimal plant growth and yield.

Water: Monitoring the presence of standing or pooled water aids in evaluating irrigation efficiency, drainage effectiveness, and waterlogging risks, guiding strategies to improve water management and prevent crop diseases.

Waterway: Analyzing waterway patterns assists in assessing drainage systems, maintaining proper water flow, preventing flooding, and supporting optimal field conditions for sustained crop growth.

Weed Cluster: Detecting concentrated weed patches facilitates targeted weed control strategies, minimizing weed interference and allowing for site-specific herbicide applications or manual removal, reducing yield losses and herbicide usage.

Endrow: Analyzing the characteristics of endrows assists in precision agriculture by identifying boundary effects Endrows provide access for machinery and equipment to enter and exit the field without damaging the crops. They allow for easier turning and maneuvering of large agricultural vehicles, such as tractors, during various field operations.

The dataset's size and shape variations further enhance its utility for agricultural research. With 94,986 images collected from 3,432 farmlands across the US, the Agriculture-Vision dataset captures the diversity and heterogeneity of real-world agricultural settings. This variability presents researchers with the opportunity to develop robust and generalizable models capable of handling different farmland layouts, cropping systems, and geographical locations. To ensure the quality and reliability of the annotations, a strict quality assurance process has been implemented during the annotation phase. This process, coupled with the expertise of agronomists, guarantees the accuracy and precision of the provided annotations, making the Agriculture-Vision dataset a trustworthy resource for semantic segmentation tasks in aerial agricultural images. [25]

Overall, the Agriculture-Vision dataset fills a critical gap in the availability of high-resolution aerial farmland images with comprehensive annotations. It opens up new avenues for research and innovation in the field of computer vision for agriculture, empowering researchers to develop advanced algorithms and tools for automated analysis and monitoring of agricultural patterns, leading to improved crop management practices and enhanced productivity in the agricultural sector.

6.4.2. Dataset preparation and pipelining

Like we see in the last section the intricate challenges associated with handling a substantial dataset, encompassing 94,986 images alongside their corresponding 8 label masks, culminating in an aggregate of approximately 190,000 images. Furthermore, the limitations imposed by hardware resources necessitate careful consideration in order to devise optimal solutions. Consequently, this study underscores the critical importance of data preprocessing, a facet that is often overlooked despite its significant impact on the overall model performance.

This section presents a comprehensive approach to address the data handling challenges associated with semantic segmentation datasets. Firstly, Proper organization of the dataset enables to efficiently navigate and access the required image and label mask pairs for training and evaluation purposes. Secondly, advanced preprocessing techniques specifically tailored for semantic segmentation are explored. These techniques encompass data augmentation methods such as random cropping, rotation, scaling, and flipping, which enhance the dataset's diversity and improve the model's ability to generalize to different viewpoints and object scales. The hardware considerations in the context of semantic segmentation are also addressed in this section. Leveraging parallel processing techniques and optimizing memory usage during data preprocessing can significantly accelerate the segmentation pipeline.

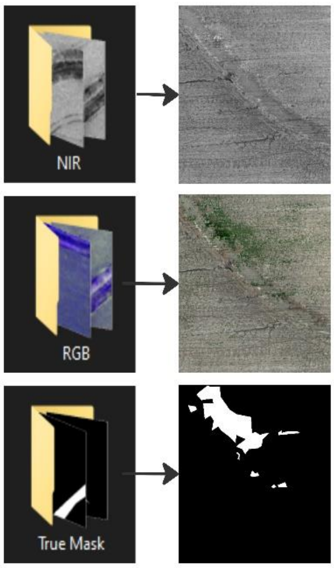
6.4.2.1. Data Preprocessing

Data preprocessing is a crucial step in preparing the dataset for semantic segmentation research. We follow the following steps in our method.

i. **Data Organization:** Organize the dataset into appropriate directories or folders, separating the images and their corresponding annotation masks. This organization facilitates easy access and management of the data during the preprocessing phase.

ii. **Data Exploration and Visualization**: Gain familiarity with the dataset by visually inspecting a sample of images and their corresponding annotation masks. This step helps in understanding the characteristics of the data and identifying any potential issues or anomalies.

iii. **Data Cleaning**: Perform any necessary cleaning steps to remove irrelevant or corrupted data from the dataset. This may involve removing duplicate images, handling missing annotation masks, or addressing any other data quality issues that may impact the training process.



CH 6 Figure 21: Data Preprocessing

iv. **Data Augmentation:** Apply data augmentation techniques to increase the diversity and variability of the dataset. Common augmentation methods for semantic segmentation include random cropping, rotation, scaling, and flipping of both the images and their corresponding annotation masks. These techniques help in enriching the dataset and improving the generalization capability of the models.

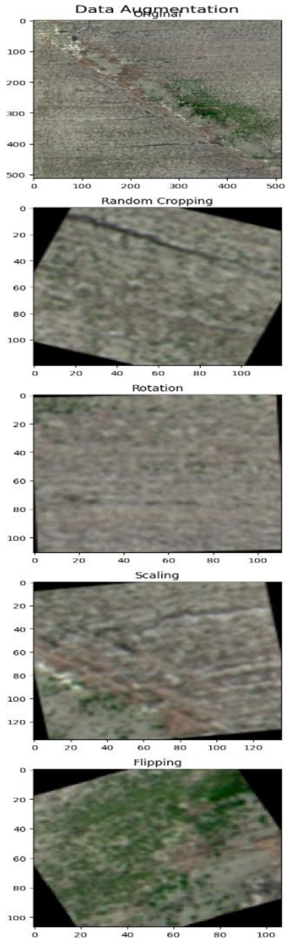
v. **Normalization:** Normalize the input images to ensure consistent and standardized pixel values across the dataset. Common normalization techniques include scaling the pixel values to a specific range (e.g., [0, 1]) or applying mean subtraction and standard deviation normalization.

vi. **Image resizing:** is one of the essential operations in deep learning, especially while using transfer learning. In this process, we take a pretrained model and modify it according to our task. So, for this we need to resize the dataset in accordance with the pretrained model. This is where image resizing comes into play. [18].

vii. **Splitting the Dataset:** Divide the dataset into separate subsets for training, validation, and testing. The training set is used to train the model, the validation set is used for tuning hyperparameters and monitoring the model's performance, and the testing set is reserved for final evaluation.

6.4.2.2. Data pipelining

As a machine learning project grows, the importance of data pipelines becomes more significant. They facilitate easy infrastructure scaling when dealing with large datasets or resource requirements. Data pipelining refers to the efficient movement of data through various stages of a machine learning pipeline, including data loading, preprocessing, and model training or inference. [26]
The input pipeline of machine learning training follows a three-stage ETL (extract, transform, and load) process. This process involves both offline and online preprocessing of raw input data before it is ready to be used for model training. [26][27]



CH 6 Figure 22: Data Augmentation

- Extract: In the extract stage, raw input data in various formats such as images, audio, or text files is collected from different sources. This can include scraping data from websites, gathering data from databases, or retrieving data from storage systems. The goal is to gather the necessary data for training the machine learning model. [26][27][28]
- Transform: The transform stage involves performing various preprocessing steps on the extracted data to make it suitable for model training. This includes tasks such as feature extraction, data cleaning, normalization, and data augmentation. For example, in the case of images, the transform stage may involve resizing images, applying data augmentation techniques like rotation, or flipping, and converting them to a standardized format. [26][27][28]
- Load: Once the data has been transformed, it is loaded into the training pipeline. This typically involves converting the preprocessed data into a binary format that can be efficiently ingested by the model during training. The binary format allows for faster loading and processing of the data, enabling efficient training on large datasets. The loaded data is then fed into the model for training. [26][27][28].

The input pipeline is responsible for fetching the raw input data from storage and transforming it into input features for the model. [26]

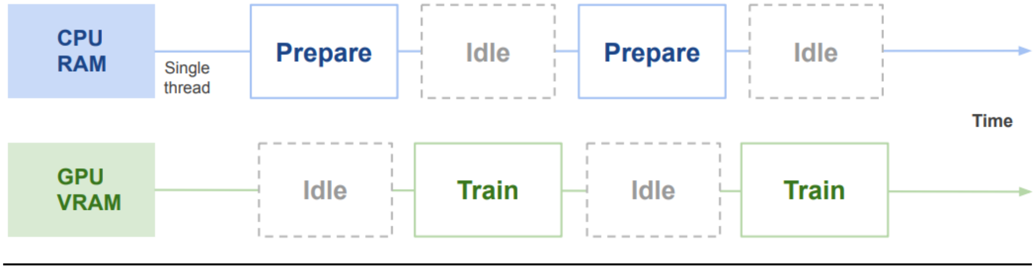
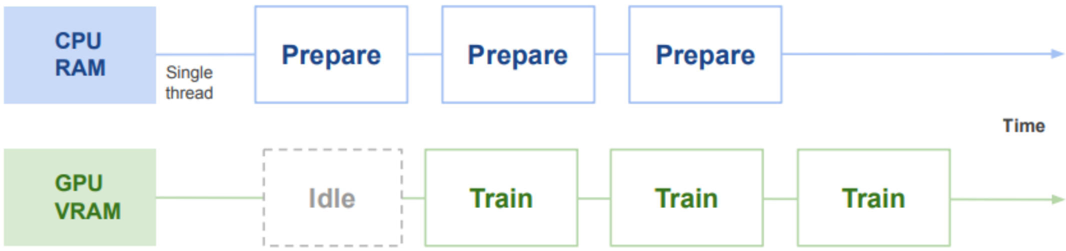


Figure 6-23: Single threaded CPU and single GPU working sequentially with no prefetching [28]

In this scenario, the CPU and GPU operate sequentially without prefetching. The CPU handles data loading and preprocessing tasks one by one, and then passes the processed data to the GPU for model training. This sequential process can result in idle time for the GPU, as it needs to wait for the CPU to finish its tasks before it can start training. As a result, the overall training process may be slower and less efficient.



CH 6 Figure 24: Single threaded CPU and single GPU working with prefetching [28]

In this scenario, prefetching is implemented to improve the training process. The CPU starts loading and preprocessing the next batch of data while the GPU is still training on the current batch. This overlap in operations allows the GPU to continuously receive new batches of data without waiting for the CPU to complete preprocessing. As a result, the GPU's idle time is minimized, leading to faster and more efficient training.

Prefetching helps reduce the latency between CPU and GPU operations by overlapping their tasks. It optimizes the data pipeline by allowing the CPU and GPU to work in parallel, utilizing the available resources more effectively. This can result in significant speed-ups, especially when working with large datasets or complex models. To build efficient pipelines we can use **TensorFlow Data.**

To build efficient pipelines, one effective tool available is TensorFlow Data. TensorFlow Data provides a high-level API that simplifies the process of creating data pipelines in TensorFlow. It offers a range of features for data manipulation, transformation, and batching, making it easier to handle complex data preprocessing tasks. By leveraging TensorFlow Data, you can streamline the input pipeline, optimize data loading and preprocessing operations, and enable parallel processing. This helps to maximize resource utilization and accelerate the training process. TensorFlow Data also integrates seamlessly with other TensorFlow components, allowing for a seamless end-to-end workflow.

By implementing robust data pipelining techniques and utilizing tools like TensorFlow Data, you can enhance the overall efficiency and performance of your machine learning workflows, ultimately leading to more accurate and reliable models.

6.4.3. Semantic Segmentation metrics

Evaluation metrics are important for assessing the performance of semantic segmentation models, which associate class labels with pixels to identify meaningful patterns in images. The choice of metric should consider the specific application of the model. Graph partitioning and superpixel-based approaches are among the methods used for evaluating segmentation models. Metrics enable fair comparison with other existing methods and play a crucial role in model validation.

6.4.3.1. Pixel Accuracy (PA):

Pixel accuracy is a simple evaluation metric for semantic segmentation models. It calculates the ratio of the total number of correctly classified pixels to the total number of pixels in the image. If there are k foreground classes and an additional background class (k+1), then the pixel accuracy is determined by dividing the sum of correctly classified pixels for all classes by the total number of pixels in the image. [29][30]

$$\mathrm{PA} = \frac{\sum_{i=0}^{K} Pii}{\sum_{i=0}^{K}\sum_{j=0}^{K} Pij}$$

6.4.3.2. Mean Pixel Accuracy (MPA)

This is an improved version of Pixel Accuracy wherein; the ratio of correctly classified pixels is calculated on the class basis. This is then averaged over the total number of classes found [29][30]

$$\mathrm{MPA} = \frac{1}{k+1}\frac{\sum_{i=0}^{K} Pii}{\sum_{i=0}^{K}\sum_{j=0}^{K} Pij}$$

6.4.3.3. Intersection over Union (IoU)

Intersection over Union (IoU) is a popular evaluation metric for semantic segmentation tasks, also known as the Jaccard Index. IoU measures the overlap between the predicted segmentation and the ground truth segmentation by computing the ratio of the area of intersection between the two sets to the area of their union. In mathematical terms, given ground truth segmentation A and predicted segmentation B, IoU is calculated as [29][30]

$$IoU = J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

TP, FP, FN - The IoU of a prediction target mask pair, if it exceeds a predefined threshold, is observed to have true positive. If the prediction mask has no associated ground truth, then a false positive is indicated. If the ground truth has no associated prediction mask, then a false negative is indicated.

$$IoU = J(A,B) = \frac{TP}{TP + FP + FN}$$

6.4.3.4. Mean Intersection over Union (MIoU)

Is a widely used evaluation metric for semantic segmentation. It calculates the IoU for each class separately and then takes the average over all classes. MIoU is expressed as the ratio of the sum of true positives for all classes to the sum of true positives, false negatives, and false positives for all classes. [29][30]

$$IoU = J(A,B) = \frac{1}{c}\sum_{c} \frac{TP}{TP + FP + FN}$$

6.4.3.5. Precision

Precision or positive predictive value (PPV), is the relation between true positives and all positive predictions: [29][30]

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

6.4.3.6. Recall

It indicates the completeness of the positive prediction to that of the ground truth. It determines of all of ground truth annotations, how many positive predictions are [29][30]

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

6.4.3.7. F1 score

This is the harmonic mean between precision and recall. It brings in a balance between precision and recall. A good F1 score implies less false positives and less false negatives [29][30]

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

6.4.3.8. Dice coefficient

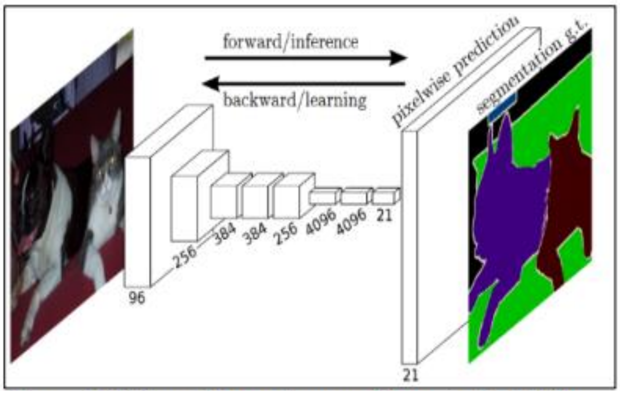
This is defined as two times the overlap area between the prediction and ground truth map, divided by the sum of pixels in both prediction and ground truth map. Suppose A is the ground truth and B is the predicted segmentation truth, then [29][30]

$$Dice = 2\,\frac{|A \cap B|}{|A| + |B|}$$

In summary, evaluation metrics play a crucial role in assessing the performance of semantic segmentation models and enabling fair comparison with existing methods. The choice of metric should consider the specific application of the model.

6.4.4. Models

6.4.4.1. Fully Convolutional Networks (FCN)



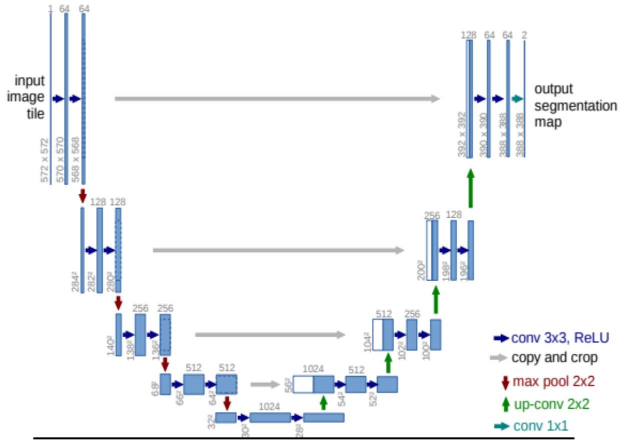
CH 6 Figure 25: FCN workflow diagram (Long et al., 2015)

Fully Convolutional Neural Networks (FCNN) has become a popular strategy and baseline for semantic segmentation in recent years. FCNNs replace fully connected layers with convolutional layers, allowing them to handle input images of arbitrary sizes and produce corresponding outputs with spatial dimensions. Chen et al. (2014) introduced a similar FCNN model but integrated conditional random fields (CRFs) for detailed boundary recovery. Long et al. (2015) [31] proposed an architecture that combined semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations. The workflow of FCNNs is illustrated in Figure 6-25, where each layer of data in a Convnet is a three-dimensional array of size h x w x d. Convolutional networks are built on translation invariance and operate on local input regions with relative spatial coordinates. FCNNs can efficiently compute feedforward computation and back propagation over an entire image, making them a popular choice for semantic segmentation tasks. [31][32]

6.4.4.2. U-Net

The U-Net is a convolutional neural network architecture that was introduced in 2015 as a solution for biomedical image segmentation tasks. It has gained popularity due to its impressive performance and has been applied to a variety of other domains such as satellite and aerial image analysis, self-driving cars, and even natural disaster detection.

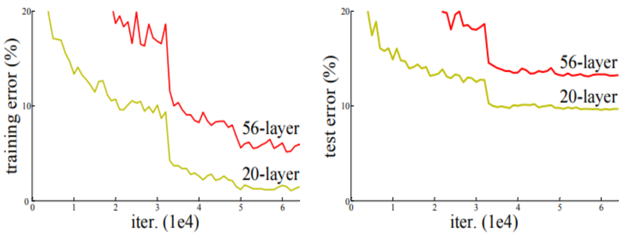
Fig. 6-26. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations. [33]



CH 6 Figure 26: U-net architecture

6.4.4.3. Residual U-Net:

Going deeper would improve the performance of a multi-layer neural network, however, could hamper the training, and a degradation problem maybe occur [4]. To overcome these problems, He et al. [4] proposed the residual neural network to facilitate training and address the degradation problem. The residual neural network consists of a series of stacked residual units. Each residual unit can be illustrated as a general



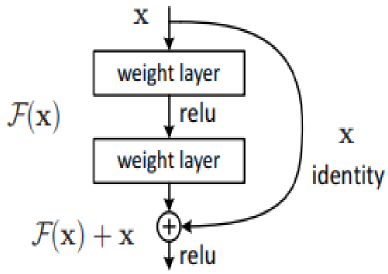
CH 6 Figure 27: Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer "plain" networks. The deeper network has higher training error, and thus tests error.

6.4.4.3.1. Residual Block

Recent research has shown that increasing the depth of a multi-layer neural network can lead to improved performance, but also runs the risk of hampering training and causing a degradation problem [34]. In particular, it has been observed that as network depth increases, accuracy can become saturated and then rapidly degrade, even in the absence of over fitting. [34] [35]

To address this problem, He et al [35] proposed the residual neural network, which consists of a series of stacked residual units. Each residual unit allows the network to learn a residual mapping of the input, rather than directly fitting the desired underlying mapping. This approach has been shown to be effective in improving performance and overcoming the degradation problem [34]. Experimental results have demonstrated the effectiveness of the residual neural network approach. For example, He et al [34] thoroughly verified the degradation problem and the effectiveness of the residual neural network approach through experiments on the CIFAR-10 and ImageNet datasets.

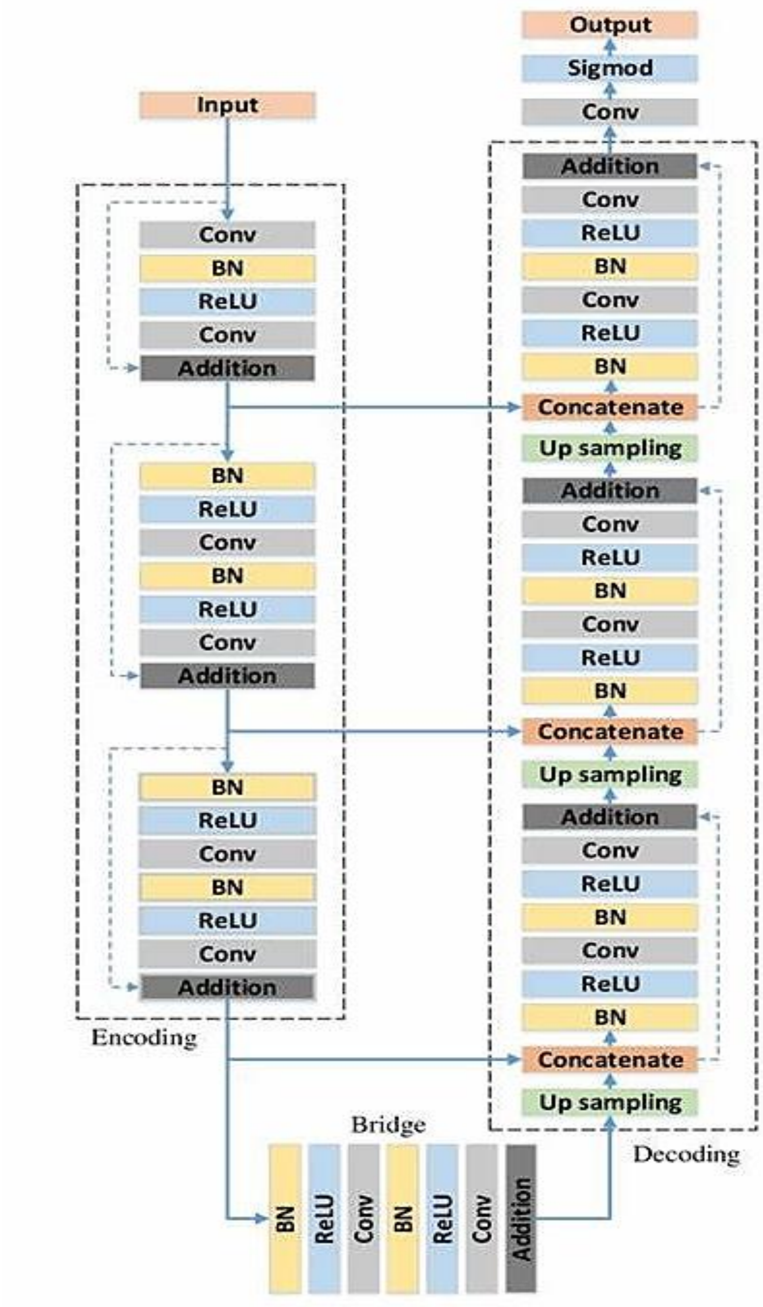
Huang et al [34] similarly observed the degradation problem in experiments on deep residual networks and demonstrated that the use of residual connections can alleviate the problem. These results suggest that the residual neural network approach is a promising approach for improving the performance of deep neural networks. [34]



CH 6 Figure 28: Residual learning: a building block.

Residual U-net (ResUnet)

Deep ResUnet is a semantic segmentation neural network that combines the strengths of both U-Net and residual neural network architectures. This combination offers two benefits: First, the residual unit helps to ease the training of the network. Second, the skip connections within a residual unit and between low and high levels of the network facilitate information propagation without degradation, enabling the design of a neural network with significantly fewer parameters while still achieving comparable or even better performance on semantic segmentation tasks. [35]



CH 6 Figure 29: Residual U-net

6.4.4.4. U-net ++

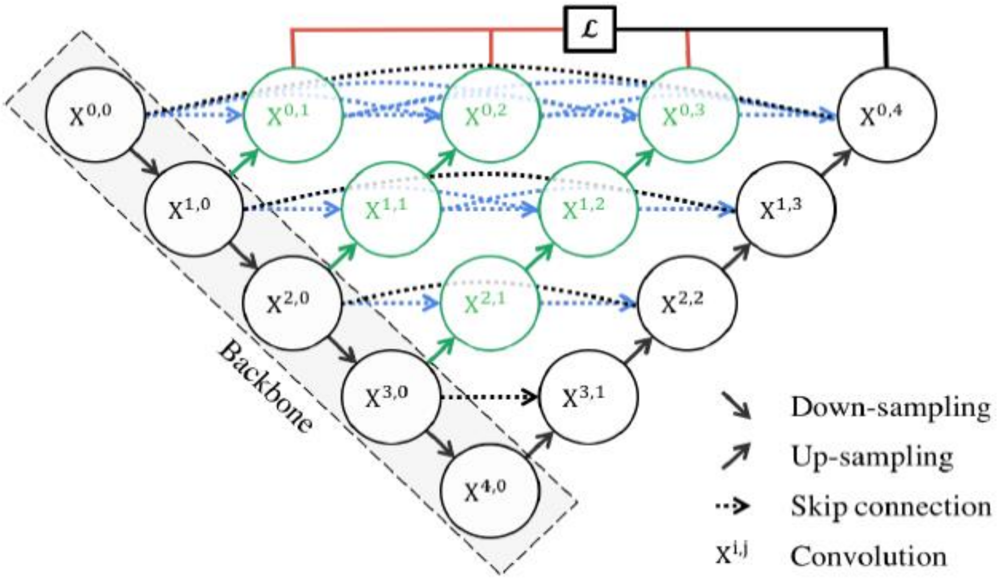
The UNet++ architecture is a novel segmentation approach that improves upon the original U-Net model. It achieves this by incorporating three key enhancements: convolution layers on skip pathways, dense skip connections, and deep supervision. [36]

i. **The convolution layers on skip pathways**, which bridge the semantic gap between encoder and decoder feature maps. [36]
ii. **The dense skip connections**, which improve gradient flow, allowing for more effective information propagation. [36]
iii. **The inclusion of deep supervision**, which enables model pruning and improves or in the worst case achieves comparable performance to using only one loss layer. [36]

6.4.4.4.1. UNet++ Architecture

UNet++ consists of an encoder and decoder that are connected through a series of nested dense convolutional blocks. The main idea behind UNet++ is to bridge the semantic gap between the feature maps of the encoder and decoder prior to fusion. For example, the semantic gap between (X0, 0, X1,3) is bridged using a dense convolution block with three convolution layers. In the graphical abstract, black indicates the original U-Net, green and blue show dense convolution blocks on the skip pathways, and red indicates deep supervision. Red, green, and blue components distinguish UNet++ from U-Net. [36]

Overall, UNet++ is a powerful segmentation architecture that has proven to be effective in reducing semantic gaps and improving segmentation accuracy. Its re-designed skip pathways and deep supervision make it a promising approach for addressing segmentation challenges at multiple scales.

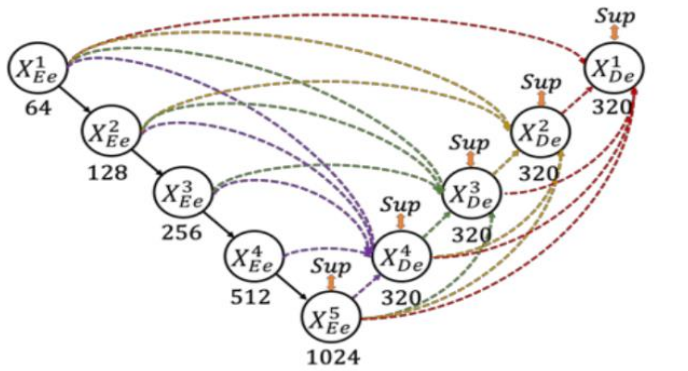


CH 6 Figure 30: UNet++ Architecture

6.4.4.5. U-Net 3+

A new architecture called U-Net 3+ was proposed by Oktay et al. in 2020. UNet 3+ utilizes full-scale skip connections and deep supervisions to extract more information from all scales. This architecture was developed by researchers from Zhejiang University, Sir Run Run Shaw Hospital, Ritsumeikan University, and Zhejiang Lab. [37]

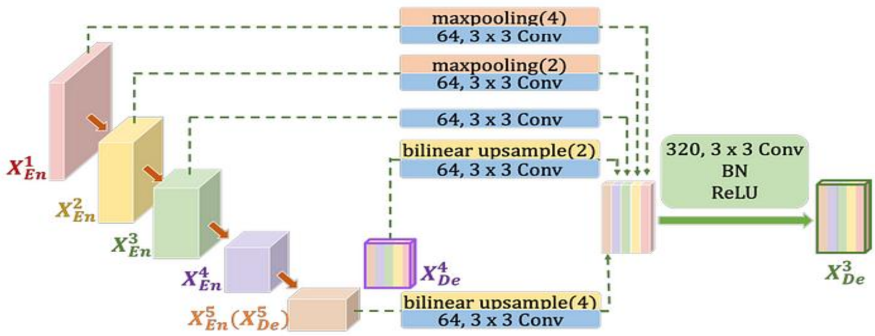
i. **Full-scale skip connections:** incorporate low-level details with high-level semantics from feature maps in different scales.
ii. **Full-scale deep supervision:** learns hierarchical representations from the full-scale aggregated feature maps.



CH 6 Figure 31: Full-scale deep supervision

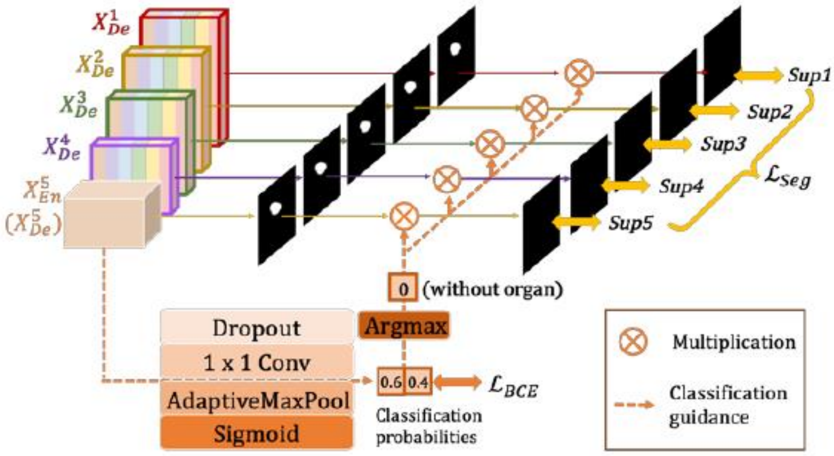
6.4.4.5.1. Full-scale skip connections

The UNet 3+ architecture addresses the limitations of previous models, such as UNet and UNet++, by incorporating full-scale skip connections. These connections capture fine-grained details and coarse-grained semantics in full scales, enabling the model to learn the position and boundary of an organ more explicitly. In UNet 3+, each decoder layer integrates both smaller- and same-scale feature maps from the encoder and larger-scale feature maps from the decoder. This is achieved through a set of inter encoder-decode skip connections that deliver low-level detailed information from the smaller-scale encoder layers and a chain of intra decoder skip connections that transmit high-level semantic information from larger-scale decoder layers. This approach helps to capture a wider range of information and improve the accuracy of biomedical image segmentation. [37]



CH 6 Figure 32: Full-scale skip connections

6.4.4.5.2. Full-scale Deep Supervision



CH 6 Figure 33: Full-scale Deep Supervision

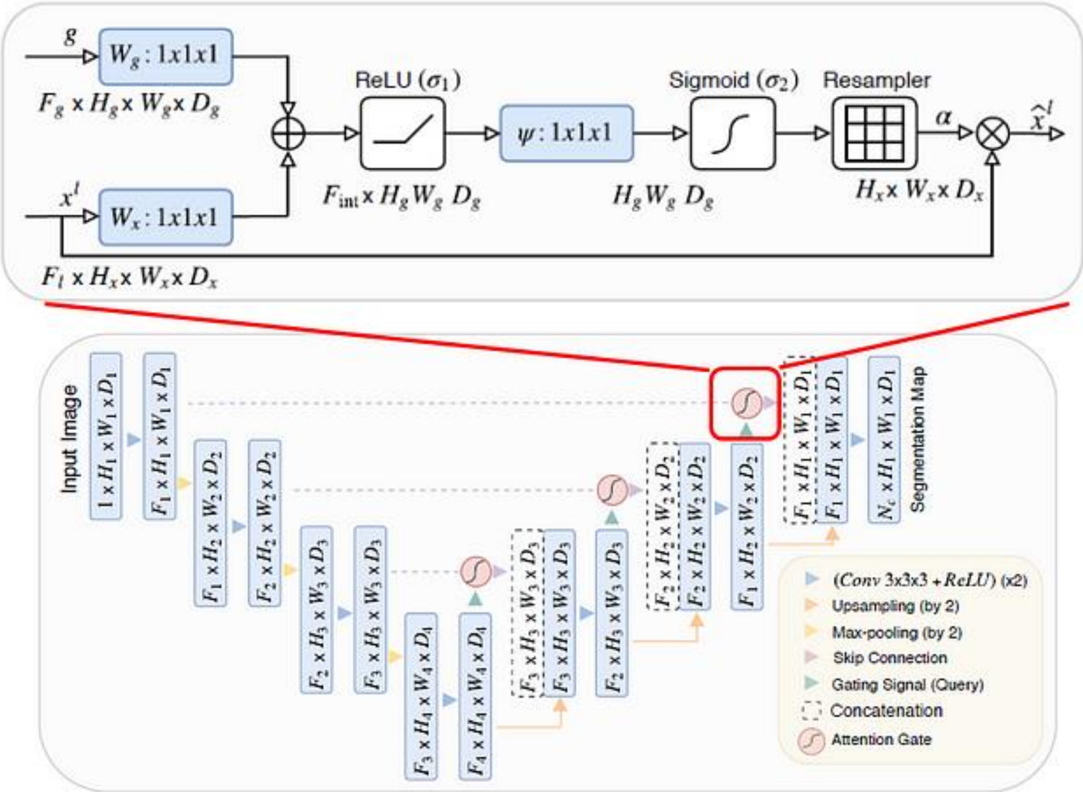
UNet 3+ architecture incorporates a deep supervision mechanism through the utilization of side outputs from each decoder stage (Sup1 to Sup5) that are supervised by the ground truth. To achieve this, the last layer of each decoder stage is fed into a plain 3 × 3 convolution layer, which is then followed by a bilinear up-sampling operation and a sigmoid function. This allows for the generation of multiple intermediate segmentation outputs, thereby promoting more efficient and effective training of the network. [37]

Summarize the differences between Unet and Unet ++ and Unet +++

The U-Net is popular convolutional neural network architecture for biomedical image segmentation. It was proposed by Ronneberger et al. in 2015 and has been widely used in various medical imaging applications. To improve the performance of U-Net, a new architecture called UNet++ was proposed by Zhou et al. in 2018. UNet++ uses nested and dense skip connections to bridge the semantic gap between encoder and decoder feature maps, as well as deep supervision to enable model pruning and improve performance. However, UNet++ does not explore sufficient information from full scales. To address this limitation, a new architecture called UNet 3+ was proposed by Oktay et al. in 2020, which uses full-scale skip connections and deep supervisions to further improve performance.

6.4.4.6. Attention U-Net

Attention U-Net, by Imperial College London, Nagoya University & Aichi Cancer Center, University of Luebeck, HeartFlow, and Babylon Health, is briefly reviewed. [38] With Attention Gate (AG), the model automatically focuses to learn the target structures of varying shapes and sizes. [38]



CH 6 Figure 34: Attention U-Net

6.4.4.6.1. Architecture of Attention U-Net

The architecture of Attention U-Net is similar to the U-Net and it has a contraction path and an expansion path. The contraction path includes a series of convolutional layers and max pooling operations, which extracts the local features from the input image. On the other hand, the expansion path consists of upsampling layers followed by convolutional layers, which capture the global features of the image.

One unique aspect of Attention U-Net is that there is an attention gate (AG) at each skip connection. The AG learns to selectively highlight the relevant regions of the feature maps from the contraction path, which are used to refine the corresponding feature maps from the expansion path. This mechanism helps the network to focus on important features and suppress the irrelevant features, which can improve the segmentation accuracy. [38]

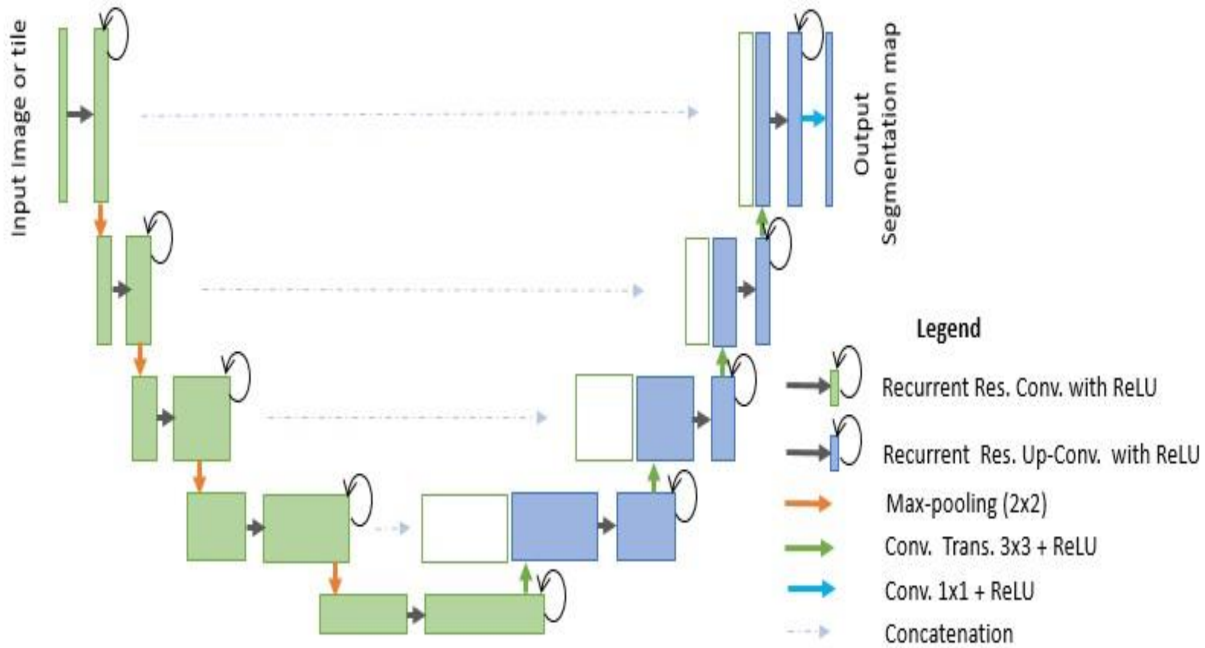
6.4.4.6.2. Attention Gate

The concept of a "Gated Attention Mechanism" (GAM), which is a neural network module that combines both attention and gating mechanisms. The GAM is used to enhance the features in a convolutional neural network (CNN) by selectively focusing on the most relevant information from different feature maps. The attention mechanism is responsible for identifying the most important features, while the gating mechanism controls the flow of information by selectively allowing or blocking the attention weights. This allows the CNN to effectively attend to the most important features and ignore irrelevant ones, leading to better performance on tasks such as image classification. [38] [39]

The gate is composed of two branches: a feature branch and a gating branch. The feature branch applies a series of convolutional operations to the input, while the gating branch learns to produce a spatial attention mask that emphasizes or suppresses certain regions of the input. The two branches are then multiplied element-wise to produce an attended feature map, which preserves the most informative features while ignoring the noise or irrelevant features. The attended feature map is then passed on to the subsequent layers of the network for further processing. [38] [39]

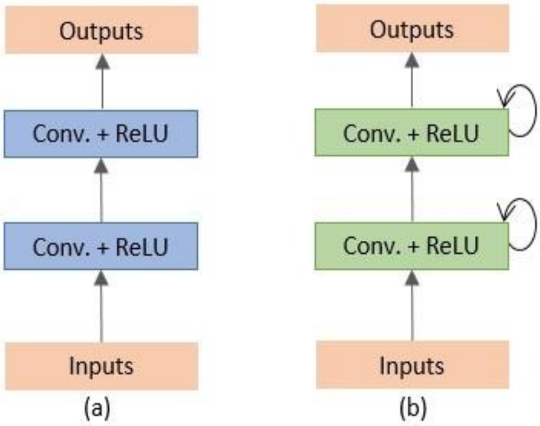
6.4.4.7. Recurrent U-Net

The Recurrent U-Net (RU-Net) is a modification of the popular U-Net architecture that incorporates recurrent connections to capture long-term dependencies in the feature maps. [39][40]



CH 6 Figure 35: Recurrent U-Net

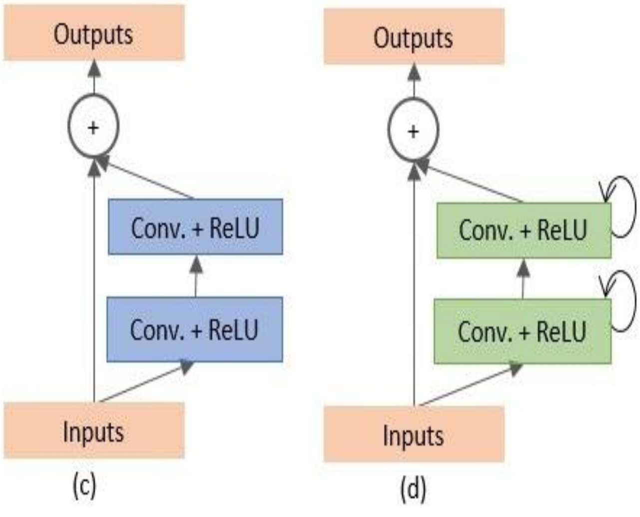
The R2U-Net architecture is composed of a contracting path and an expanding path, similar to the U-Net. However, each block in the contracting and expanding paths contains recurrent connections, which are implemented using gated recurrent units (GRUs). The use of GRUs allows the R2U-Net to selectively update the hidden state of each block based on the current input and the previous hidden state. [39][40]



CH 6 Figure 36: figure, (a) represent the normal block of U-net, (b) represent the block of RU-Net.

6.4.4.8. Recurrent Residual U-Net (R2-UNet)

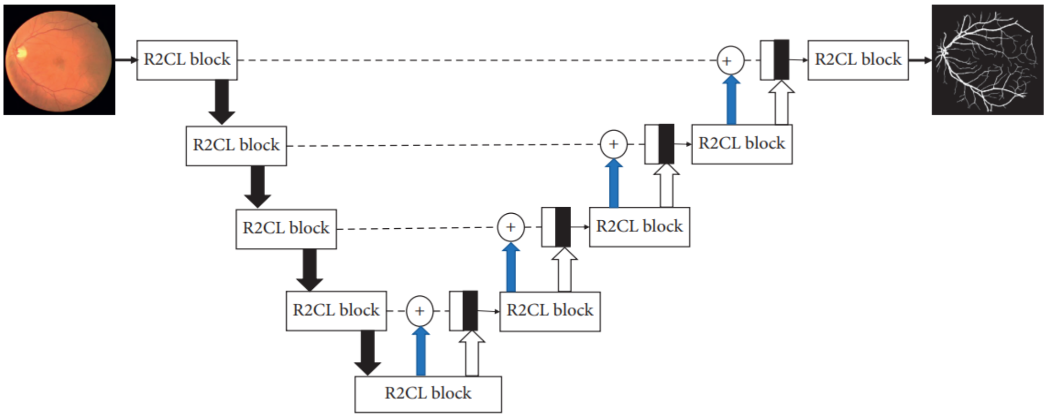
The R2-UNet adds residual connections to the R2U-Net architecture. Residual connections allow the network to learn the residual features from the input, making it easier to optimize the network and reduce the vanishing gradient problem. The R2-ResUNet is able to achieve state-of-the-art performance on medical image segmentation tasks. [39][40]



CH 6 Figure 37: figure, (c) represent the residual block, (d) represent the block of R2-UNet

6.4.4.9. Recurrent Residual U-Net with Attention gate

The Recurrent Residual U-Net with Attention gate (RRU-Net with AG) is a modification of the original U-Net architecture that incorporates recurrent and residual connections to improve feature representation and an attention gate to enhance the focus on relevant regions that introduce in 2021. [39]



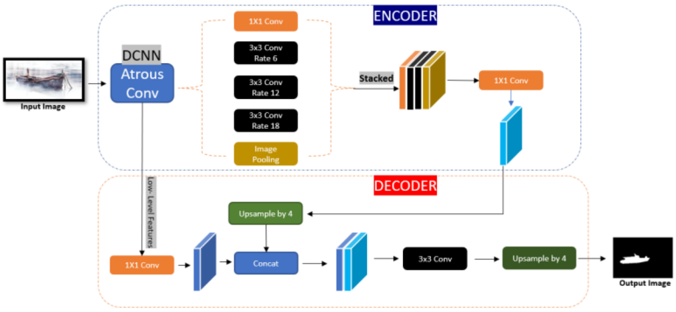
CH 6 Figure 38: Recurrent Residual U-Net with Attention gate

6.4.4.10. DeepLabv3+

DeepLabv3+ is a state-of-the-art image segmentation model that achieves highly accurate pixel-level predictions for a given image. It is an extension of the DeepLabv3 architecture, which uses atrous convolution and multi-scale processing to capture both local and global context in the input image. The architecture also employs a novel feature pyramid network (FPN) module that aggregates multi-scale feature maps to capture more diverse spatial information in the image. [41][42][43]

Figure (6-39) (deeplabV3+) [43]

The figure shows the structure of DeeplabV3+, which is composed of an encoder and decoder path with an ASPP module at the bottleneck of the structure.



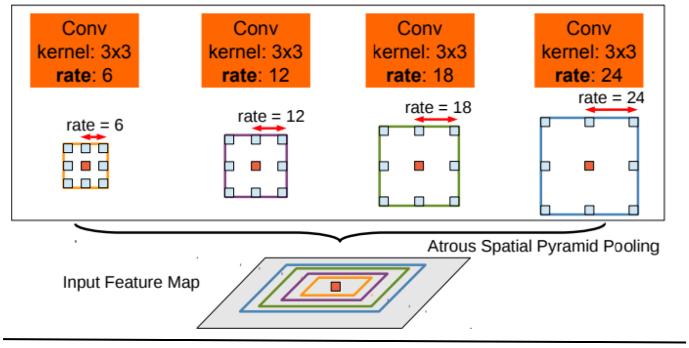
CH 6 Figure 39: the structure of DeeplabV3+

The DeepLab V3+ architecture is designed to enhance the receptive field of each convolution output without sacrificing information, by utilizing dilated convolution and an encoder-decoder structure. In the encoding stage, the main DCNN deep convolutional neural network uses serial Atrous Convolution. The output of the main network is then split into two parts, one directly fed to the decoder branch, and the other passed through a parallel Atrous Convolution layer, which employs different rates for feature extraction. The output is then compressed using 1x1 volume stitching. The decoding stage takes two inputs: the output of the main network and the result of the parallel dilation and convolution. After processing, the two results are merged and upsampled using bilinear interpolation.[41][42][43][44]

One unique aspect of the DeepLab V3+ network is the use of atrous spatial pyramid pooling (ASPP), which involves employing atrous convolution in parallel as a strategy to extract features at multiple scales and to alleviate the loss of spatial information due to pooling or convolutions with striding operations. This technique allows increasing the field of view and spatial context at each layer without increasing the number of parameters and computational complexity significantly.[41][43][44]

Figure (6-40) (ASPP) [44]

The figure shows The ASPP (Atrous Spatial Pyramid Pooling) technique in DeepLab V3+ uses multiple parallel filters with different rates to exploit multi-scale features and increase the effective field-of-view (FOV) of each layer.



CH 6 Figure 40: The ASPP (Atrous Spatial Pyramid Pooling) technique in DeepLab V3+

5.2 Results

In this section, we will delve into the results obtained from the various models mentioned in the methodology section. Firstly, we will showcase the analytical scores for each model across different datasets. Additionally, we will demonstrate the progression of evaluation metrics, such as the Jaccard coefficient, during the runtime of the models. Furthermore, we will present side-by-side comparisons of the model predictions and the corresponding ground truth data. This visual representation will provide insights into the accuracy and performance of each model in capturing the desired outcomes. Lastly, we will provide a specific example of a farm land where the models successfully detect and precisely determine the affected area and quantify the extent of damage. This example will serve as a tangible demonstration of the models' capabilities in addressing crop-related issues. By presenting these comprehensive results, we aim to provide a thorough understanding of the performance and effectiveness of the models employed in predicting and assessing crop issues.

5.3.1 Assessing Performance

The section will introduce the results of models in different resolution which will demonstrate how the selection of different types of data or labeling can significantly affect the performance of machine learning models. The evaluation metrics used in section 2, such as precision, accuracy, and Jaccard coefficient, are fundamental measures of model performance and can be used to guide the selection of data and labeling for training and evaluation.

Deep learning models have become increasingly popular in recent years due to their ability to learn complex patterns and relationships in data. However, these models can also be sensitive to the type of data and labeling used for training and evaluation. The table highlights the importance of careful consideration and selection of appropriate data and labeling for achieving optimal performance of deep learning models.

In summary, the selection of appropriate data and labeling is critical for achieving optimal performance of machine learning models, including deep learning models. The evaluation metrics used in section 2, such as precision, accuracy, and Jaccard coefficient, can help guide this selection process. Careful consideration of data and labeling can help ensure that deep learning models perform at their best, enabling researchers to extract meaningful insights from complex data.

5.3.1.1 N-RGB resolution of 20 cm per pixel

	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
WEED	U-Net	10000	555	556	59	7236	.8569	.8667	.7257	.6465	.7824	26
	U-Net++	10000	555	556	77	8288	.8562	.8864	.7275	.6609	.7929	24
	U-Net3+	10000	555	556	78	15600	.8787	.8851	.7645	.6849	.8091	20
	Attention U-Net	10000	555	556	59	7888	.8541	.8679	.7147	.6414	.7796	25
	R2U-Net	10000	555	556	40	13777	.8196	.8358	.6876	.6041	.7479	75
	DeeplabV3+	10000	555	556	46	10809	.8379	.8455	.7091	.6194	.7479	137
DRYDOWN	U-Net	15125	840	841	73	14580	.8405	.8988	.7624	.7005	.8223	26
	U-Net++	15125	840	841	57	9955	.8417	.8772	.7774	.6986	.8211	24
	U-Net3+	15125	840	841	35	16572	.8012	.8655	.7272	.6448	.7801	20
	Attention U-Net	15125	840	841	89	16868	.8962	.9256	.8403	.7851	.8789	25
	R2U-Net	15125	840	841	33	11931	.8149	.8768	.7375	.6686	.7993	75
	DeeplabV+	15125	840	841	23	8299	.8287	.8884	.7441	.6786	.8071	137
ND	U-Net	11250	625	625	104	14826	.9334	.9173	.8877	.812	.8989	26
	U-Net++	11250	625	625	52	6572	.8488	.8155	.7408	.6311	.7719	24
	U-Net3+	11250	625	625	43	13100	.7963	.7442	.6631	.5354	.6931	20
	Attention U-Net	11250	625	625	49	7186	.8091	.7777	.6785	.5657	.7213	25
	R2U-Net	11250	625	625	53	17632	.8592	.8445	.7475	.6519	.7877	75
	DeeplabV3+	11250	625	625	22	6085	.8072	.7637	.6766	.5586	.7139	137
D	U-Net	5610	311	312	46	4714	.9585	.7395	.7429	.5714	.7237	26
	U-Net++	5610	311	312	46	3669	.9556	.7177	.7435	.5666	.7215	24

OUPLE	**U-Net3+**	5610	311	312	53	11008	.9612	.7114	.7498	.5641	.7146	20
	Attention U-Net	5610	311	312	46	3269	.9667	.7904	.7286	.6001	.7477	25
	R2U-Net	5610	311	312	40	6942	.9609	.7373	.7661	.5914	.7391	75
	DeeplabV3+	5610	311	312	20	3064	.9553	.6679	.7113	.5055	.6603	137
WATER	**U-Net**	1939	107	108	31	1126	.9076	.8871	.7975	.6681	.8001	26
	U-Net++	1939	107	108	33	1081	.9251	.8815	.8381	.7181	.8351	24
	U-Net3+	1939	107	108	55	3905	.9342	.8847	.8923	.7954	.8851	20
	Attention U-Net	1939	107	108	31	995	.9143	.9005	.8081	.6907	.8162	25
	R2U-Net	1939	107	108	50	3593	.9377	.9024	.6907	.8092	.8939	75
	DeeplabV3+	1939	107	108	47	3037	.9342	.9084	.8499	.7816	.8761	137

WWAY	**U-Net**	3509	195	196	54	3596	.9814	.9331	.8645	.9271	.9545	26
	U-Net++	3509	195	196	81	4694	.9843	.9539	.8291	.8998	.9471	24
	U-Net3+	3509	195	196	33	4741	.954	.8397	.8157	.7118	.8306	20
	Attention U-Net	3509	195	196	14	718	.9598	.9071	.8391	.765	.8661	25
	R2U-Net	3509	195	196	45	5867	.9501	.8611	.8111	.7131	.8316	75
	DeeplabV3+	3509	195	196	57	4635	.9721	.9018	.8834	.7921	.8831	137
PS	**U-Net**	2339	130	131	49	1723	.9901	.837	.8691	.6738	.8031	26
	U-Net++	2339	130	131	25	923	.9696	.9067	.6215	.2901	.4361	24
	U-Net3+	2339	130	131	25	3087	.9801	.7811	.7998	.6189	.7581	20
	Attention U-Net	2339	130	131	40	1473	.9871	.8701	.8691	.6946	.8181	25
	R2U-Net	2339	130	131	46	3854	.9878	.7693	.8078	.6439	.7796	75
	DeeplabV3+	2339	130	131	22	1557	.9781	.7701	.8196	.6457	.7801	137
ENDR	**U-Net**	4033	224	225	71	4685	.9785	.9075	.9061	.8266	.9047	26
	U-Net++	4033	224	225	67	4322	.9651	.3851	.8371	.7191	.8361	24
	U-Net3+	4033	224	225	78	13590	.9736	.8681	.8872	.779	.8569	20
	Attention	4033	224	225	113	7487	.9765	.8837	.8848	.7901	.8824	25

O	**U-Net**											
W	**R2U-Net**	4033	224	225	61	8185	.9638	.8137	.8499	.7118	.8314	75
	DeeplabV 3+	4033	224	225	88	10227	.9866	.9296	.9455	.8781	.9349	137

CH 6 Table 2: compare between different models (N-RGB resolution of 20 cm per pixel).

Based on the information provided in the previous section, we can calculate the average score of each model across different labels. This approach allows us to obtain an overview of the best-performing models.
Interestingly, the results show that the U-Net model consistently achieves the highest score compared to the other models. However, if we use the best model specifically for each label, we can further enhance the results by an additional 5 percent.

This finding suggests that by selecting the most suitable model for each specific crop issue or label, we can improve the overall performance of the system. It highlights the importance of model selection and customization based on the specific characteristics and requirements of different agricultural problems.

	model	mIOU
1	best model	77.45375
2	**U-net**	72.89
3	**Attention Unet**	68.51
4	**R2U-net**	68.4125
5	**U-net+**	62.45375
6	**Deeplab V3+**	61.495
7	**U-net +++**	59.28375

CH 6 Table 3: Jaccard coefficient average score over models (res 20 cm)

5.3.1.2 N-RGB resolution of 40 cm per pixel

	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
WEED	**U-Net**	10000	555	556	92	975	.9451	.9292	.8974	.8353	.9102	26
	R2U-Net	10000	555	556	50	869	.8409	.8406	.7185	.6188	.7645	20
	U-Net3+	10000	555	556	62	2168	.9256	.8964	.8708	.7883	.8813	75
DR	**U-Net**	15125	840	841	74	1393	.9591	.9604	.9431	.9207	.9488	26

YDOWN	**R2U-Net**	15125	840	841	72	1582	.9041	.9102	.8614	.7933	.8847	20
	U-Net3+	15125	840	841	56	2295	.9354	.9409	.9068	.8571	.9231	75
ND	**U-Net**	11977	665	666	108	1248	.9642	.9478	.9409	.8874	.9403	26
	R2U-Net	11977	665	666	41	900	.8283	.8044	.7036	.5974	.7831	20
	U-Net3+	11977	665	666	93	3759	.9572	.9317	.9283	.8684	.9296	75
DOUPLE	**U-Net**	5610	311	312	69	395	.9267	.8886	.5146	.2543	.4055	26
	R2U-Net	5610	311	312	49	490	.9658	.7131	.8291	.6078	.7557	20
	U-Net3+	5610	311	312	64	1153	.9751	.7557	.8641	.6715	.8034	75
WATER	**U-Net**	1939	107	108	72	199	.9536	.9715	.8677	.7487	.8563	26
	R2U-Net	1939	107	108	46	209	.9459	.9082	.8844	.7616	.8646	20
	U-Net3+	1939	107	108	63	348	.9656	.9562	.9166	.8751	.9333	75

WWAY	**U-Net**	3509	195	196	76	385	.9884	.9646	.9547	.8649	.9274	26
	R2U-Net	3509	195	196	76	464	.9718	.8968	.9043	.8124	.8961	20
	U-Net3+	3509	195	196	51	562	.9786	.9076	.9386	.8515	.9197	75
PS	**U-Net**	2339	130	131	65	182	.9884	.9684	.8074	.3508	.5194	26
	R2U-Net	2339	130	131	52	329	.9876	.7906	.9049	.6617	.7895	20
	U-Net3+	2339	130	131	70	608	.9927	.8703	.9309	.8071	.8922	75
ENDROW	**U-Net**	4033	224	225	38	192	.8058	.9341	.3363	.2098	.3468	26
	R2U-Net	4033	224	225	44	315	.7869	.7937	.341	.2578	.4099	20
	U-Net3+	4033	224	225	48	502	.8913	.6932	.5604	.4365	.6077	75

CH 6 Table 4: compare between different models (N-RGB resolution of 40 cm per pixel).

Interestingly, like the last section the results show that the U-Net+++ model consistently achieves the highest score compared to the other models. However, if we use the best model specifically for each label, we can further enhance the results by an additional 2 percent.

	model	mIOU
1	best model	81.46125
2	U-net +++	79.90625
3	R2U-net	74.1175
4	U-net	60.30625

CH 6 Table 5: jaccord coefficient average score over models (res40 cm).

5.3.1.3 N-RGB resolution of 80cm per pixel

	Model	Train	Val	Test	Epochs	Time(s)	Accuracy	recall	Precision	JC	DC	Size (MB)
WEED	U-Net	10000	555	556	72	2720	.9186	.9224	.8467	.7778	.8751	26
	R2U-Net	10000	555	556	71	4715	.8821	.8751	.7831	.8241	.8241	20
	U-Net3+	10000	555	556	86	10860	.9402	.9359	.8941	.8411	.9135	75
DRYDOWN	U-Net	15125	840	841	94	6273	.9502	.9532	.9298	.8868	.9411	26
	R2U-Net	15125	840	841	74	5915	.9186	.9288	.8836	.8265	.9089	20
	U-Net3+	15125	840	841	64	11126	.9412	.9536	.9115	.8725	.9319	75
ND	U-Net	11977	665	666	91	3995	.9563	.9413	.9267	.8691	.9299	26
	R2U-Net	11977	665	666	62	4327	.8528	.8316	.7441	.6436	.7831	20
	U-Net3+	11977	665	666	79	10667	.9271	.8992	.8726	.7927	.8841	75
DOUPLE	U-Net	5610	311	312	54	1186	.9339	.9095	.5576	.2971	.4581	26
	R2U-Net	5610	311	312	85	2955	.9731	.7727	.8235	.6663	.7997	20
	U-Net3+	5610	311	312	46	3219	.9129	.8001	.8379	.6828	.8111	75
WAT	U-Net	1939	107	108	49	409	.9569	.9691	.8763	.7162	.8346	26
	R2U-Net	1939	107	108	53	542	.9547	.9086	.9004	.8125	.8966	20

ER	**U-Net3+**	1939	107	108	55	1109	.9613	.9575	.8983	.8617	.9257	75
WWAY	**U-Net**	3509	195	196	107	1742	.9924	.9715	.9693	.913	.9545	26
	R2U-Net	3509	195	196	52	1034	.9683	.8952	.8902	.7987	.8879	20
	U-Net3+	3509	195	196	65	2531	.9832	.9331	.9425	.8793	.9356	75
PS	**U-Net**	2339	130	131	49	461	.9635	.9649	.4726	.1139	.2046	26
	R2U-Net	2339	130	131	71	1135	.9888	.8294	.8743	.6999	.8217	20
	U-Net3+	2339	130	131	53	1704	.9906	.8563	.9101	.7509	.8548	75
ENDROW	**U-Net**	4033	224	225	45	752	.8178	.9301	.3834	.2506	.4008	26
	R2U-Net	4033	224	225	78	1617	.9555	.8264	.7879	.6578	.7936	20
	U-Net3+	4033	224	225	60	2529	.9604	.8282	.382	.7115	.8313	75

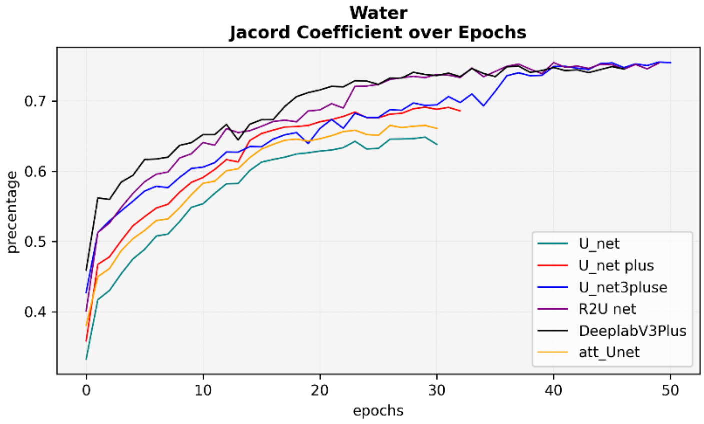
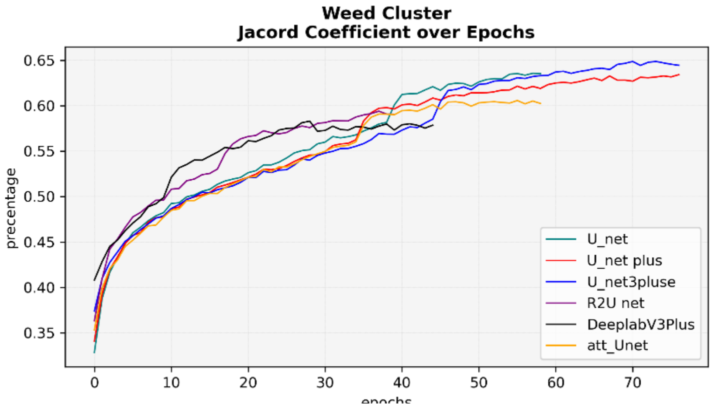
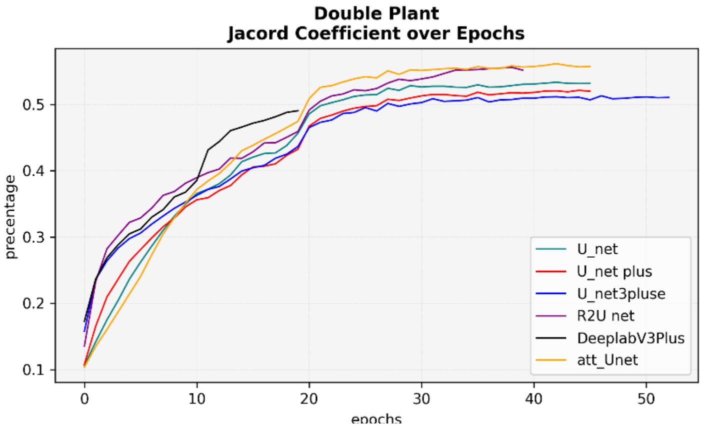
CH 6 Table 6: Compare between different models (N-RGB resolution of 80 cm per pixel).

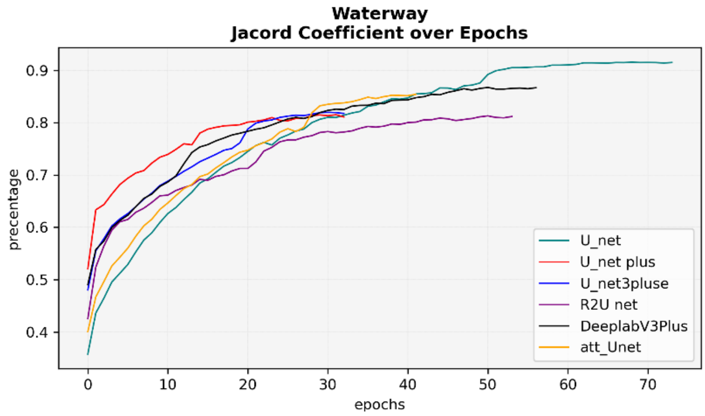
The results show that the U-Net+++ model consistently achieves the highest score compared to the other models. However, if we use the best model specifically for each label, we can further enhance the results by an additional 2 percent.

	model	mIOU
1	**best model**	78.73
2	**U-net +++**	76.93
3	**R2U-net**	63.88
4	**U-net**	63.98

CH 6 Table 7: jaccord coefficient average score over models (res80 cm).

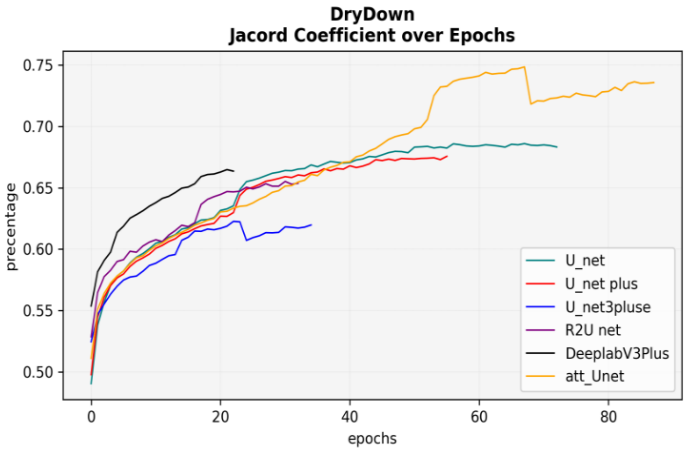
5.3.2 Comparing Model Performance during Training for Train Dataset using Jaccard Coefficient and Epochs

 CH 6 Figure 41: (Water) Jaccord coefficient over epoch	The figure compares the performance of different models during training for the water dataset. The results show that R2Unet, DeeplabV3+, and UNet3+ models have achieved similar top accuracies, indicating that they are effective in learning from the water dataset during the training phase.
 CH 6 Figure 42: (Weed cluster) jacord coefficient over epoch.	The figure compares the performance of different models during training for the Weed cluster dataset. The results show UNet3+ model has achieved the top accuracy, indicating that they are effective in learning from the Weed cluster dataset during the training phase.
 CH 6 Figure 43: (Double Plant) jacord coefficient over epoch	The figure compares the performance of different models during training for the Double Plant dataset. The results show that R2Unet, and Attention Unet models have achieved similar top accuracies, indicating that they are effective in learning from the Double Plant dataset during the training phase.



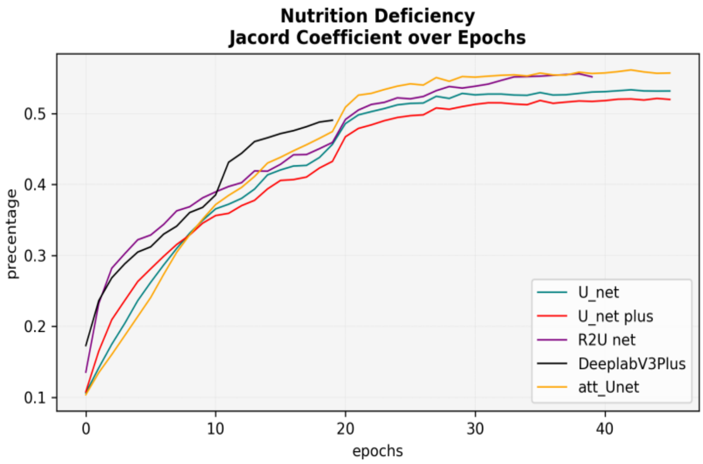
CH 6 Figure 44: (Watreway) jacord coefficient over

The figure compares the performance of different models during training for the Watreway dataset. The results show UNet model has achieved the top accuracy, indicating that they are effective in learning from the waterway dataset during the training phase



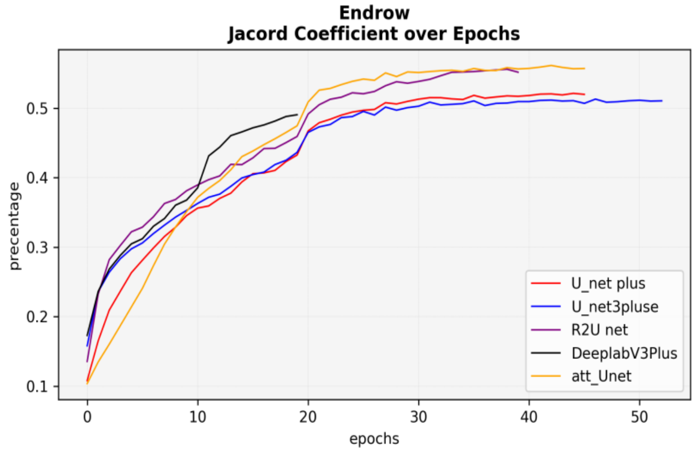
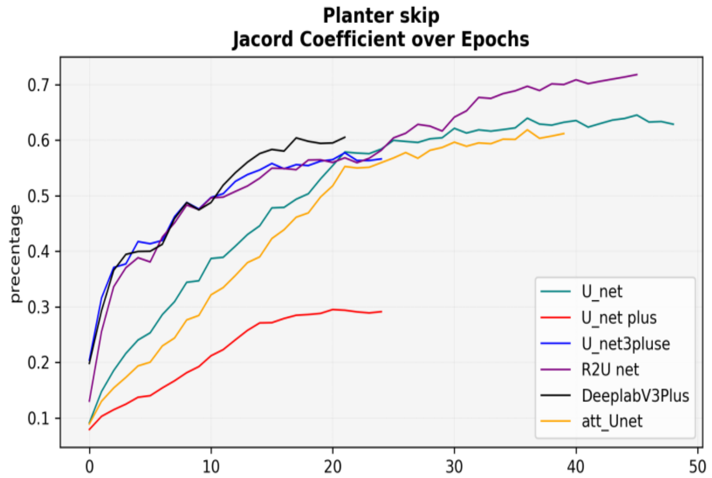
CH 6 Figure 45: (DryDowen) jacord coefficient over

The figure compares the performance of different models during training for the dry down dataset. The results show attention Unet model has achieved the top accuracy, indicating that they are effective in learning from the dry down dataset during the training phase.



CH 6 Figure 46 (Nutrition Deficiency) jacord coefficient over epoch

The figure compares the performance of different models during training for the Nutrition Deficiency dataset. The results show that deep labV3+, and Attention Unet models have achieved similar top accuracies, indicating that they are effective in learning from the Nutrition Deficiency dataset during the training phase.

 CH 6 Figure 47: (Endrow) jacord coefficient over epoch	The figure compares the performance of different models during training for the Endrow dataset. The results show that R2Unet, DeeplabV3+, and attention unet models have achieved similar top accuracies, indicating that they are effective in learning from the Endrow dataset during the training phase.
 Figure 48: (Planter skip) jacord coefficient over epoch	The figure compares the performance of different models during training for the Planter skip dataset. The results show that R2Unet model have achieved top accuracy, indicating that they are effective in learning from the Planter skip dataset during the training phase.

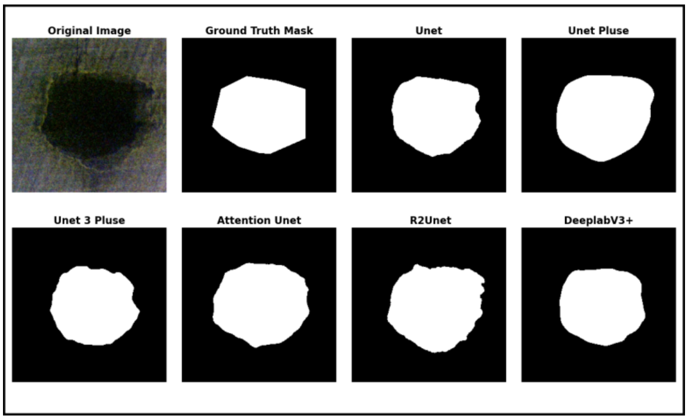
This comparison can provide insight into the effectiveness of different models and their ability to learn from the training data. A model that performs well during training, as indicated by a high JC metric, is more likely to generalize well to new, unseen data.

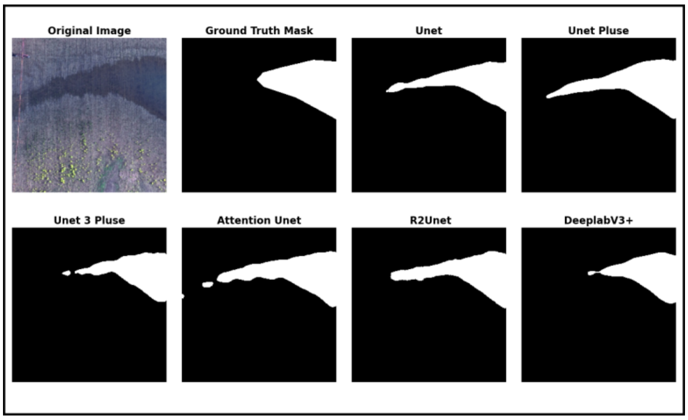
It is important to note that the optimal number of epochs for training a model may vary depending on the specific task and dataset. Therefore, it is recommended to experiment with different values for the number of epochs and other hyperparameters to find the optimal configuration for a given task.

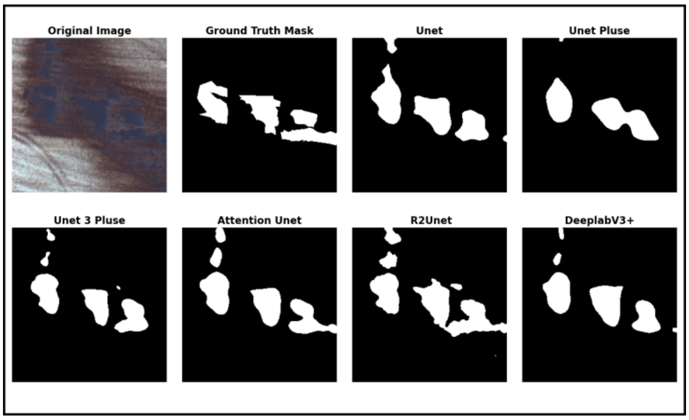
5.3.2 Comparative Study for the Prediction

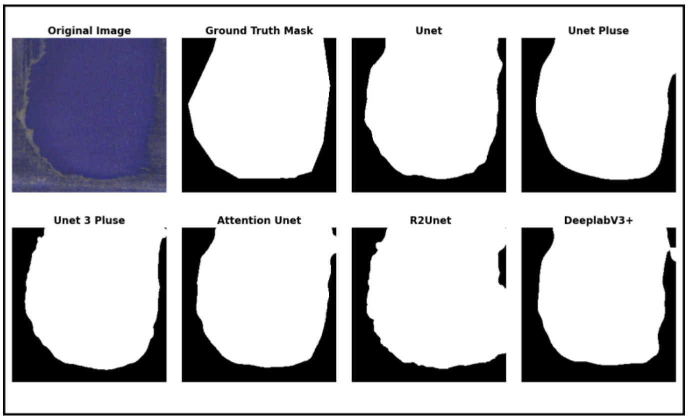
The provided images demonstrate the performance of various models, such as U-Net, U-Net++, DeepLabV3+, and others. Each model's prediction is visualized alongside the original RGB image and the ground truth mask to evaluate how effectively they can accurately segment objects or regions of interest in the images.

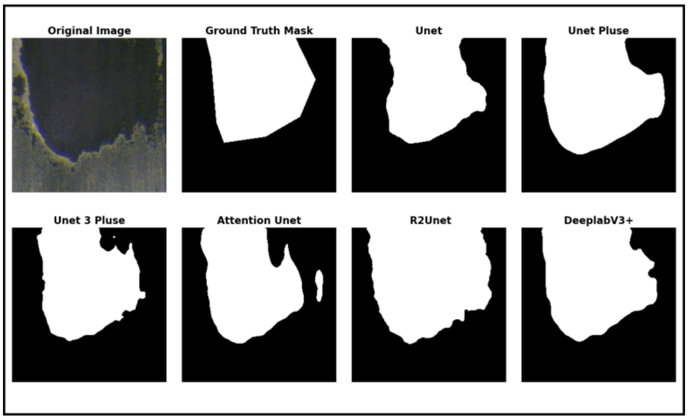
5.3.2.1 Water

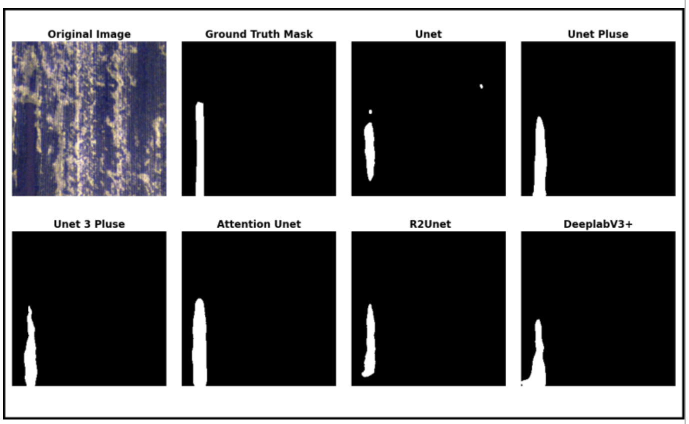




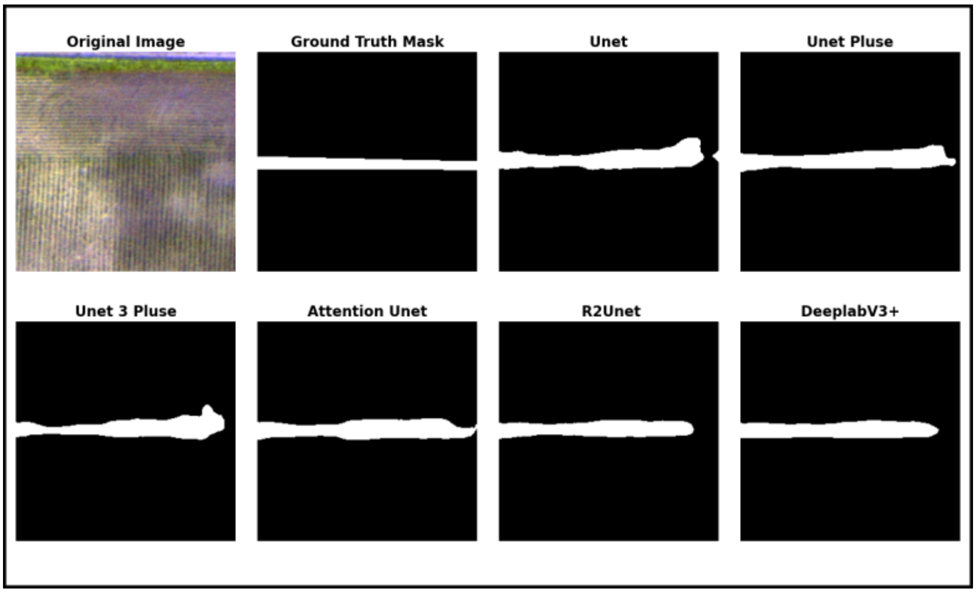


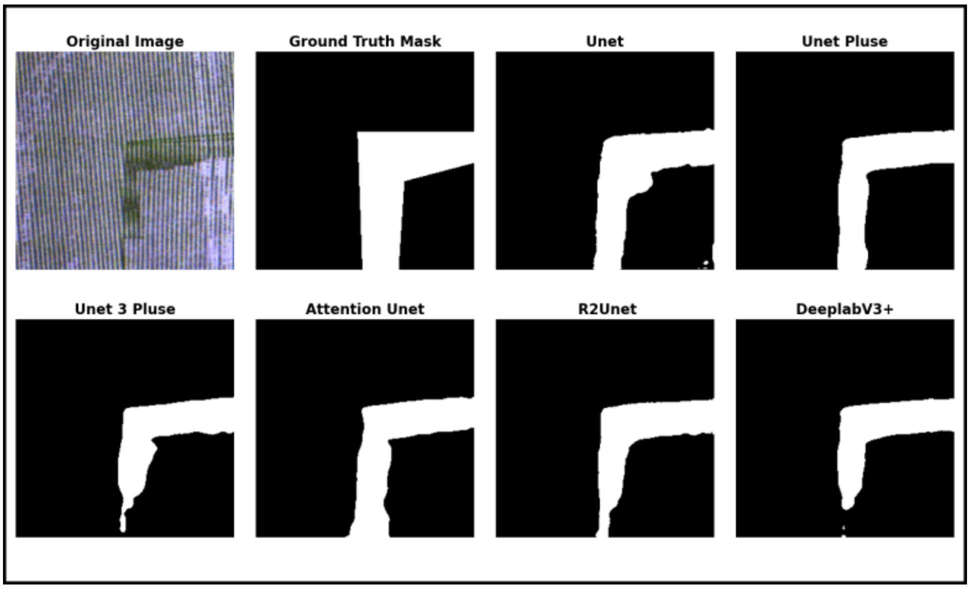


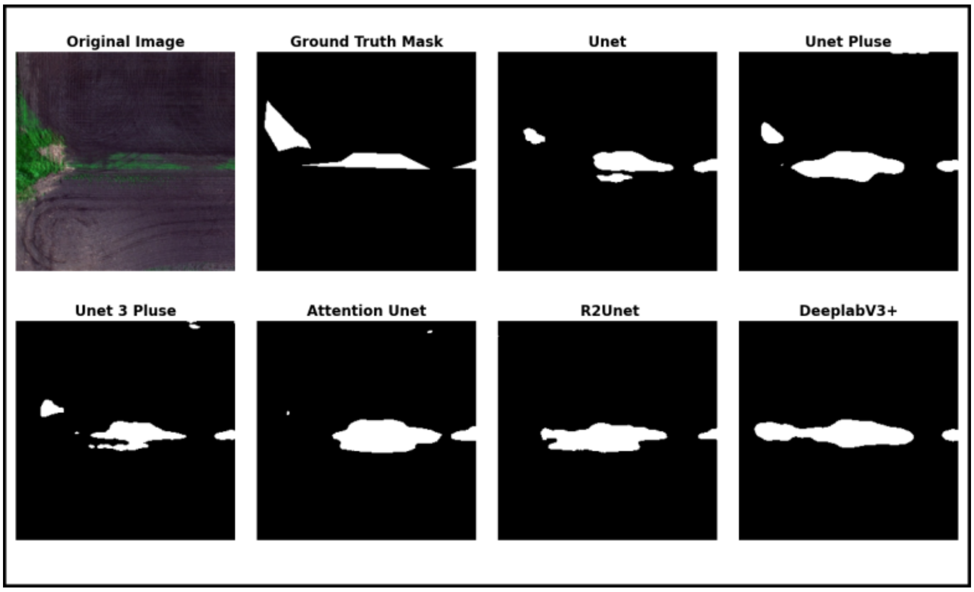




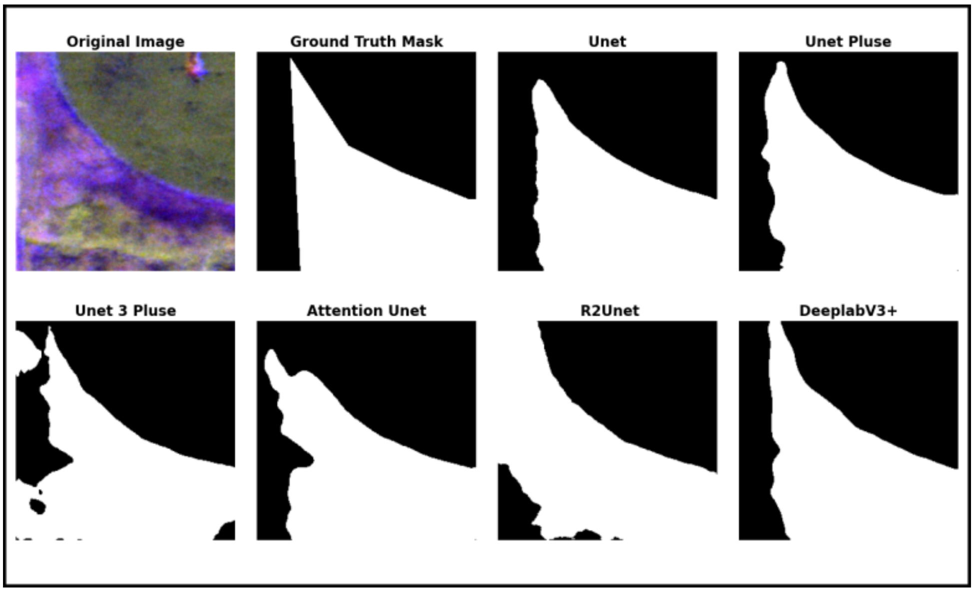
5.3.2.2 Weed_Cluster

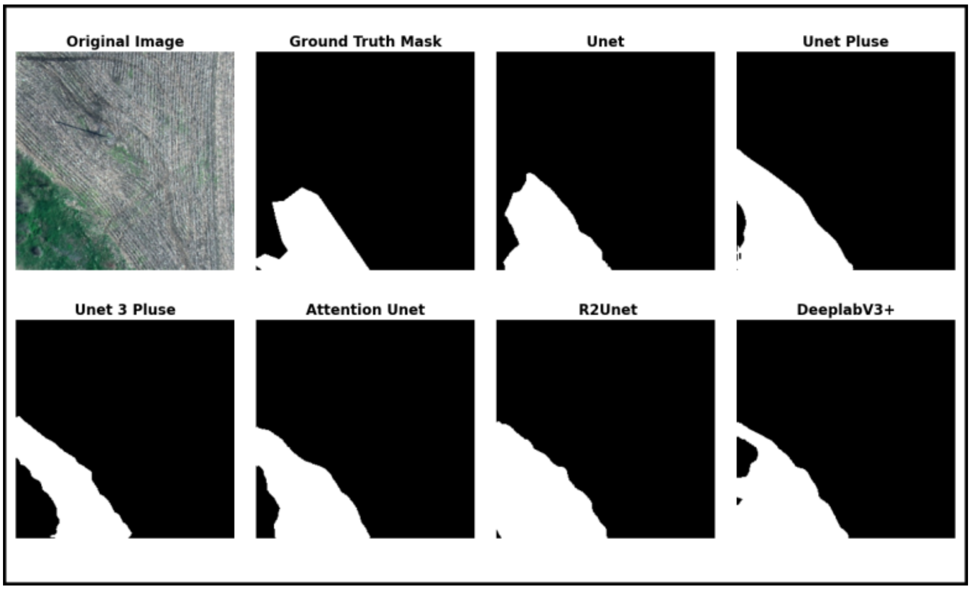

Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+

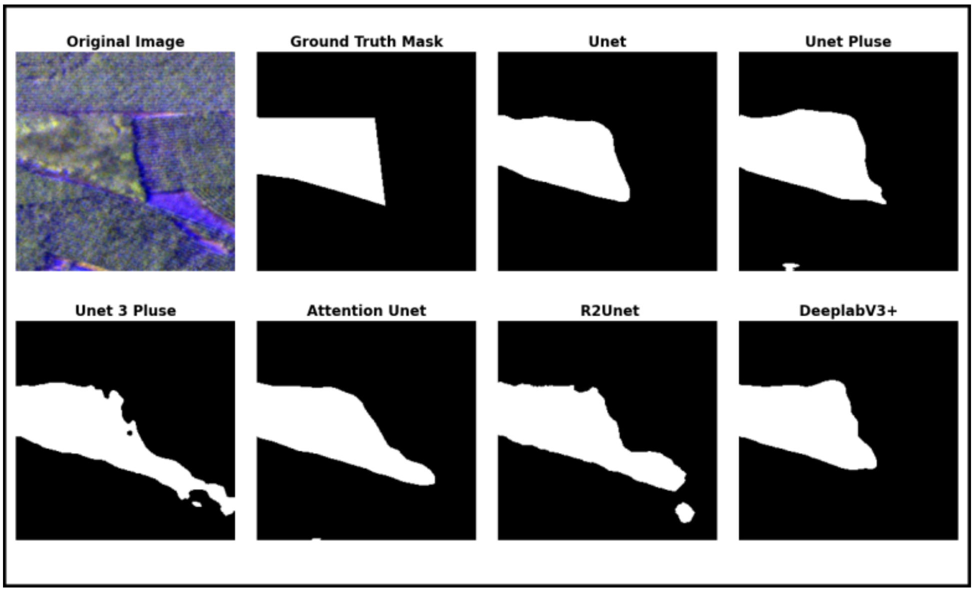

Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+


Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+

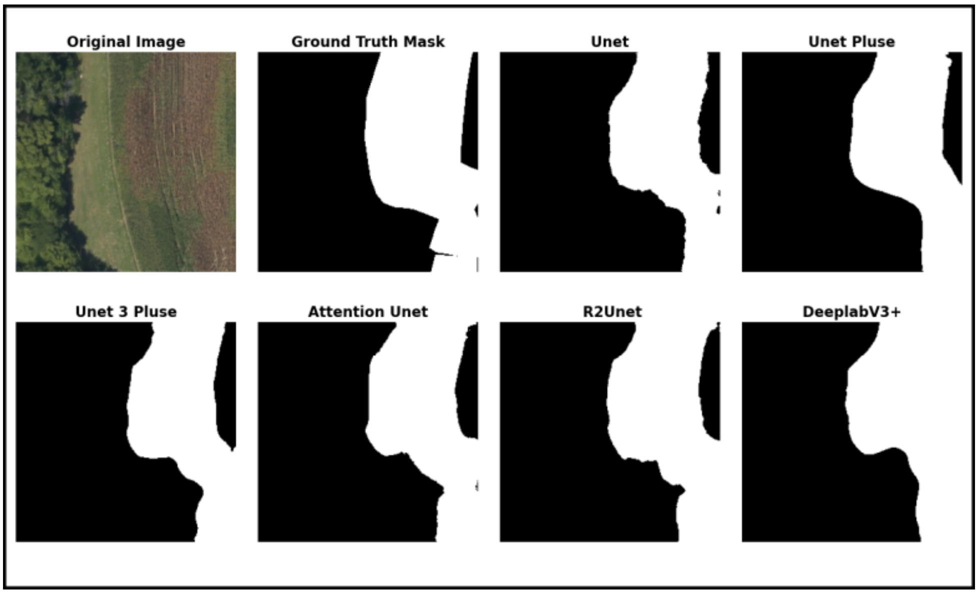
5.3.2.3 Water Way

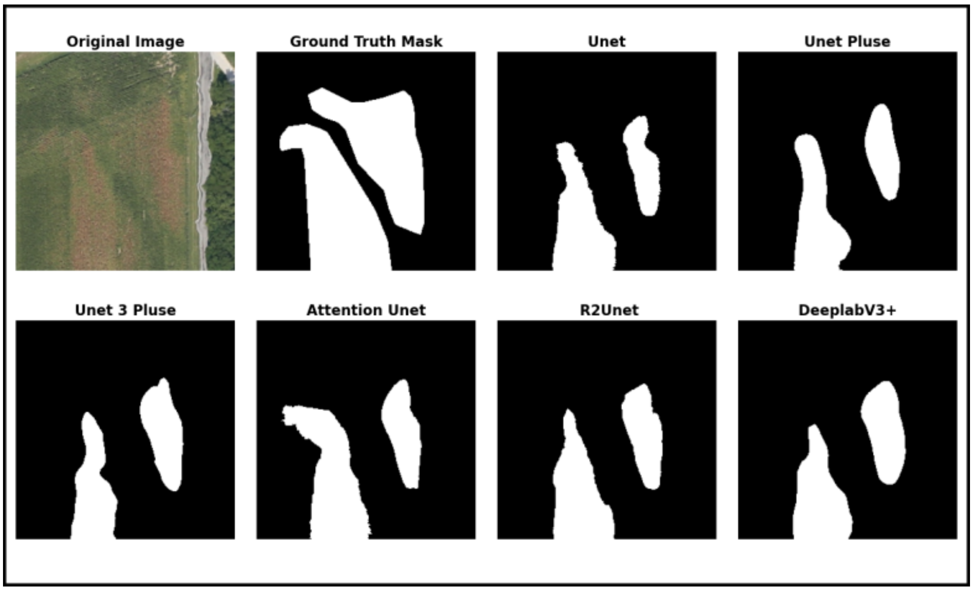


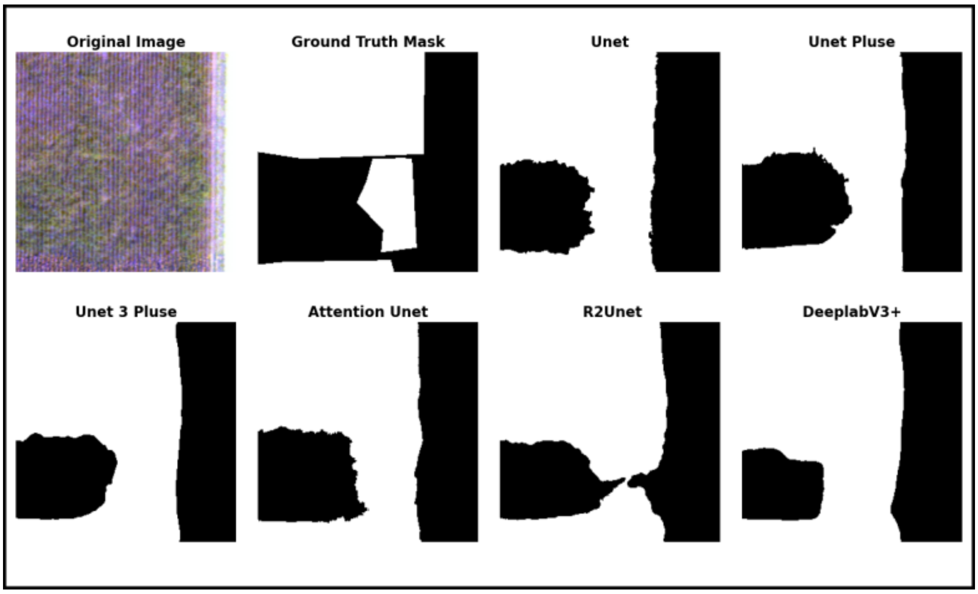




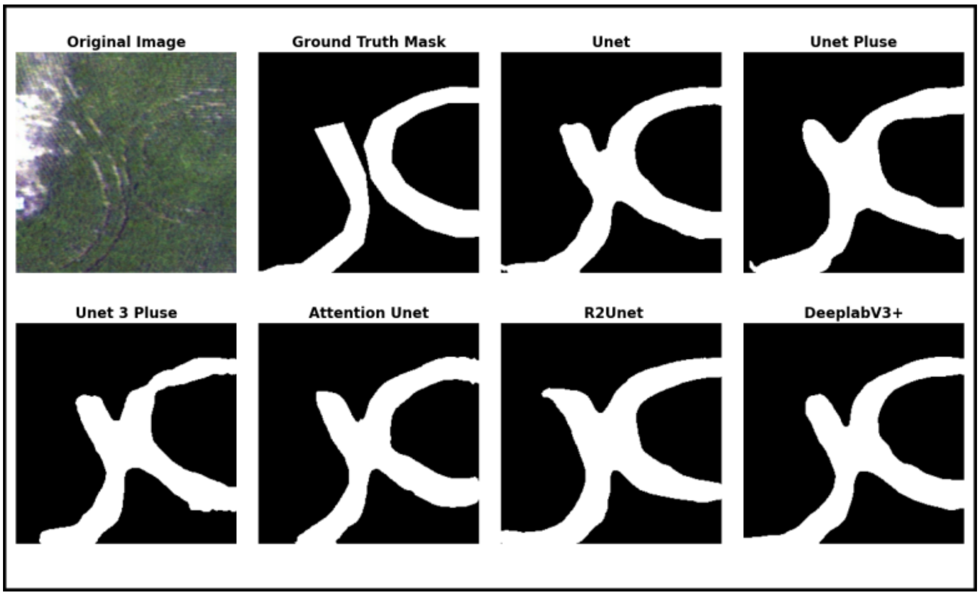
5.3.2.4 Dryness

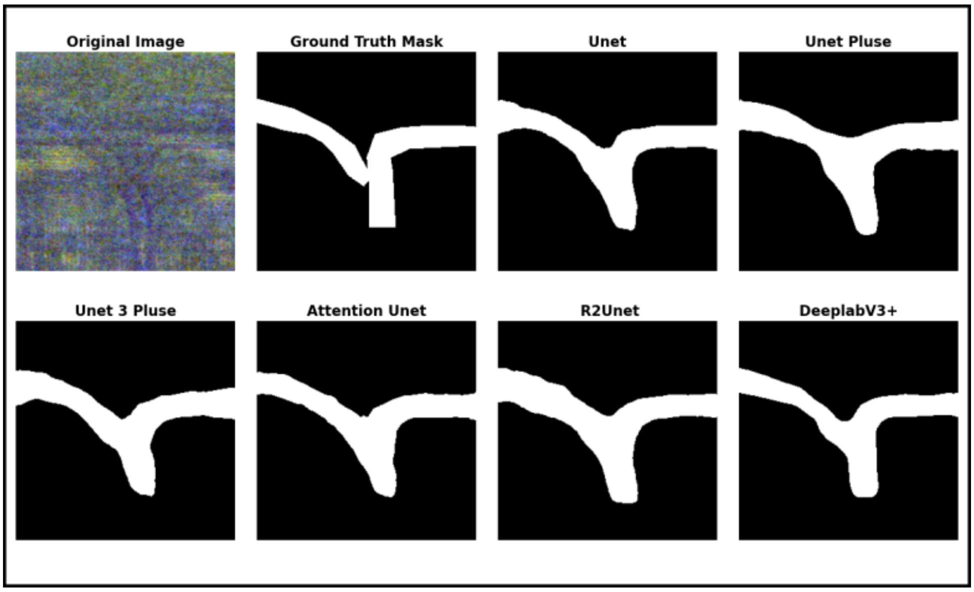


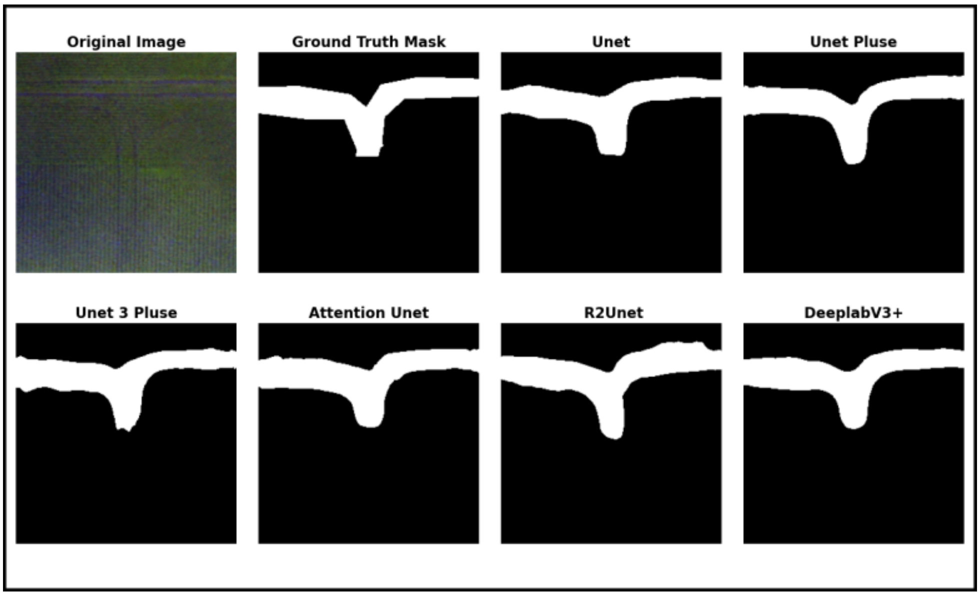




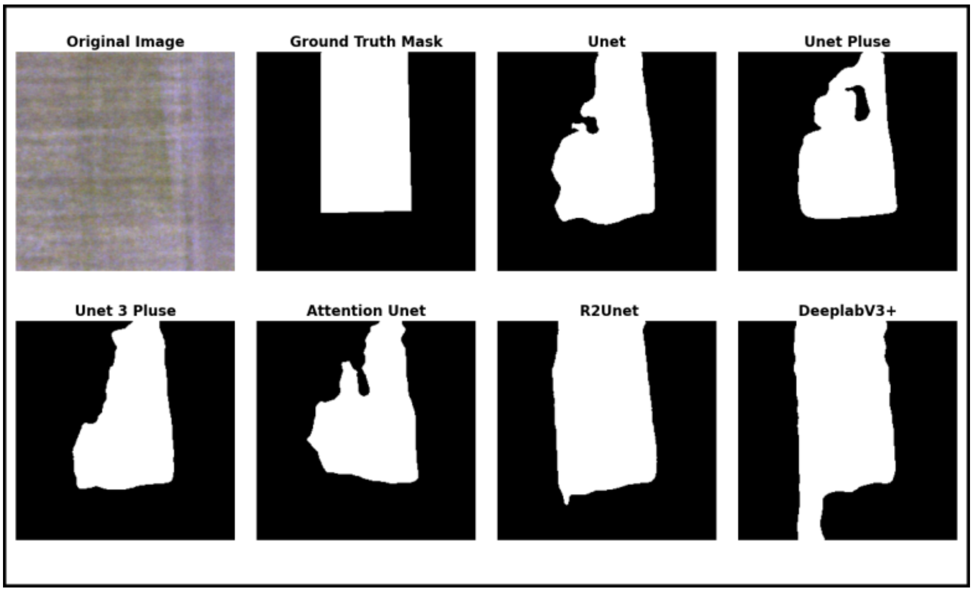
5.3.2.5 Endrow

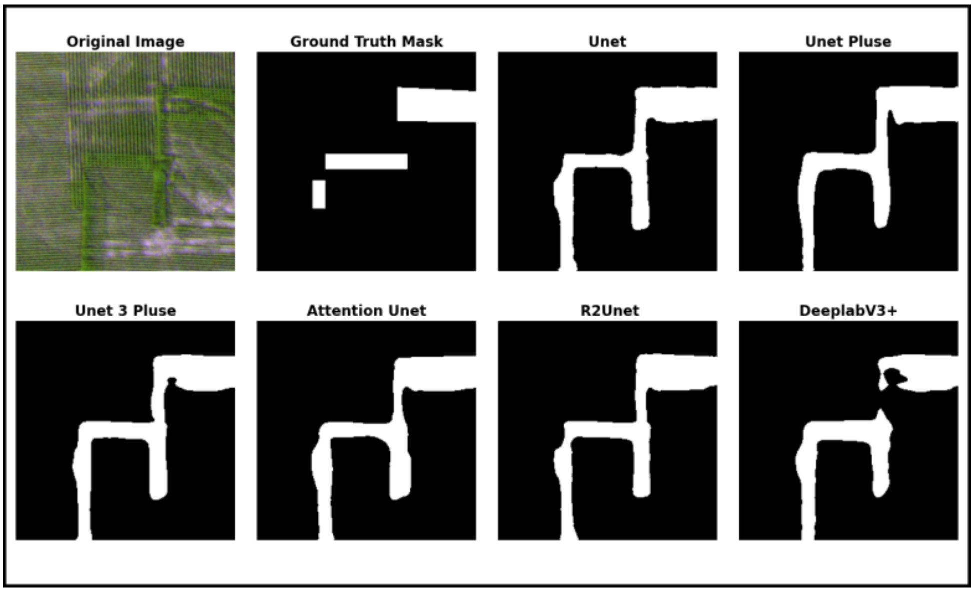

Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+

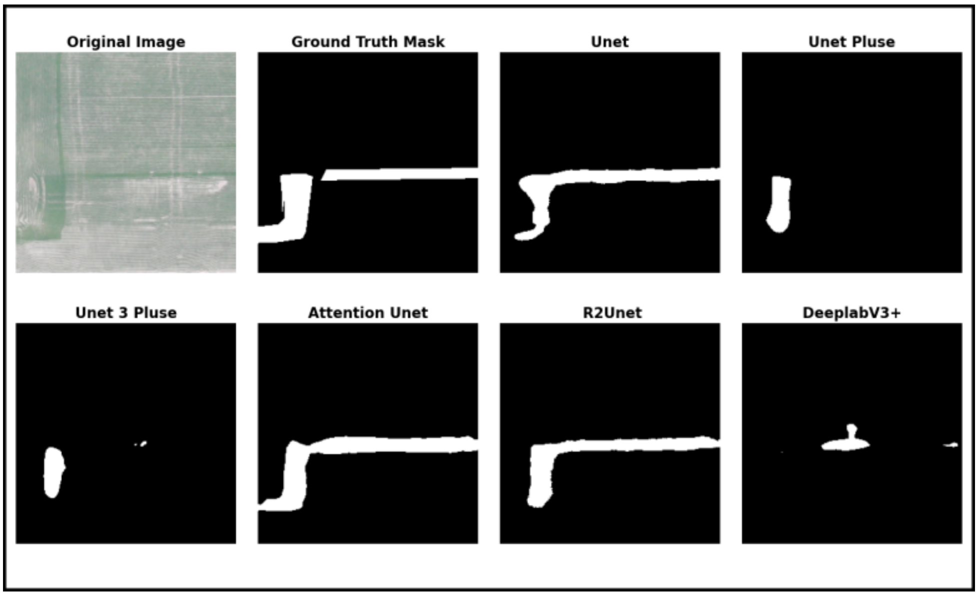

Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+


Original Image
Ground Truth Mask
Unet
Unet Pluse
Unet 3 Pluse
Attention Unet
R2Unet
DeeplabV3+

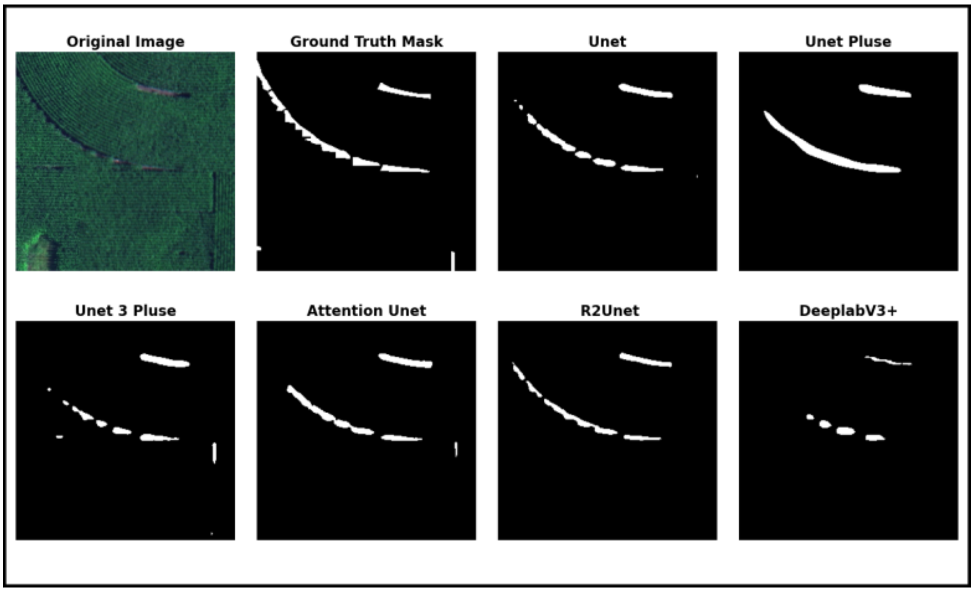
5.3.2.6 Double Plant

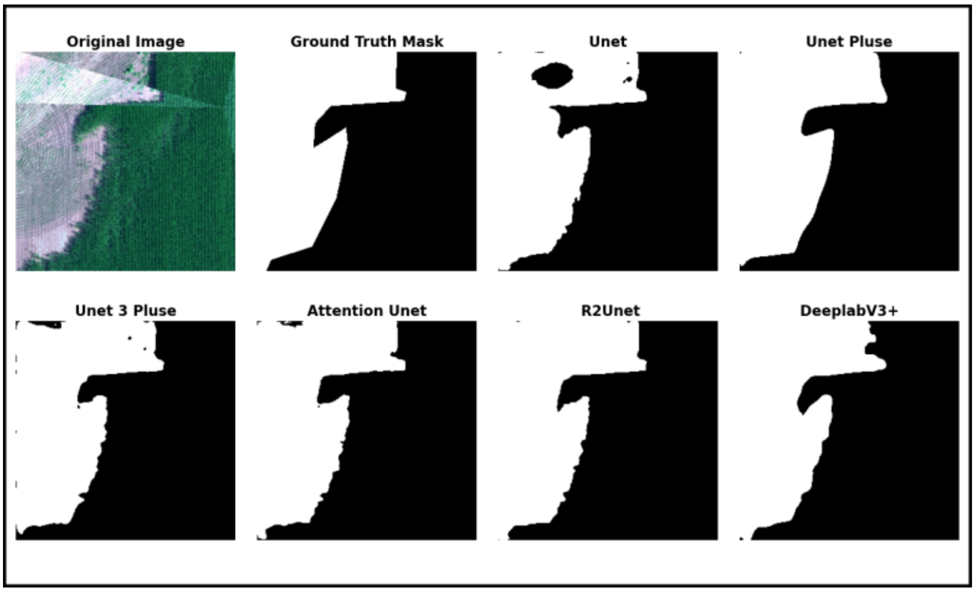


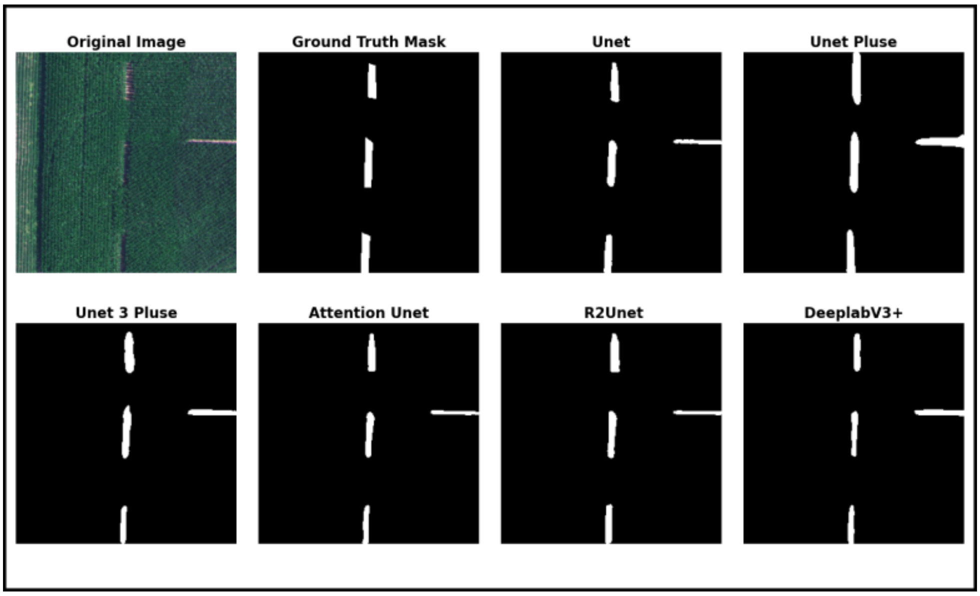




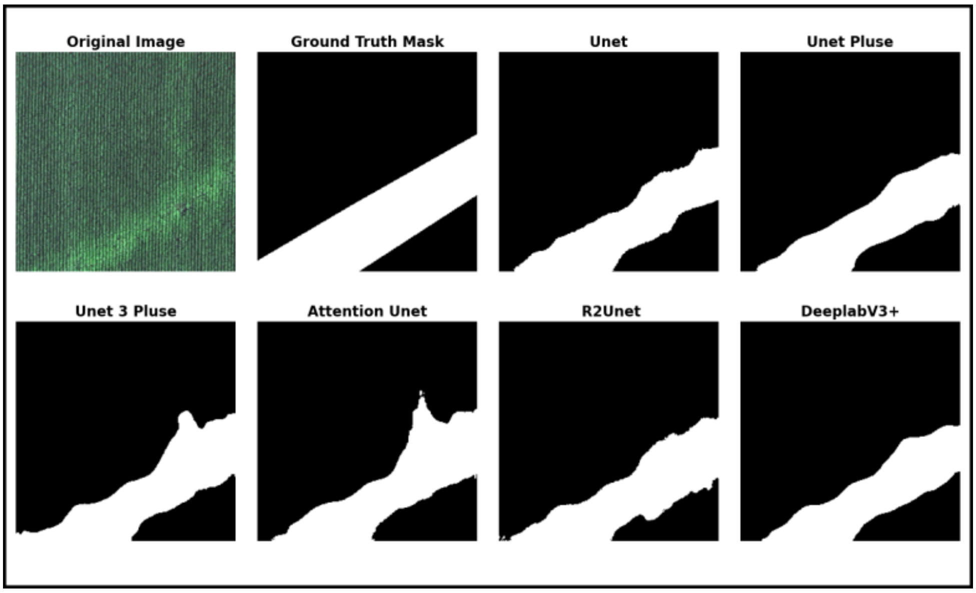
5.3.2.7 Planter Skip

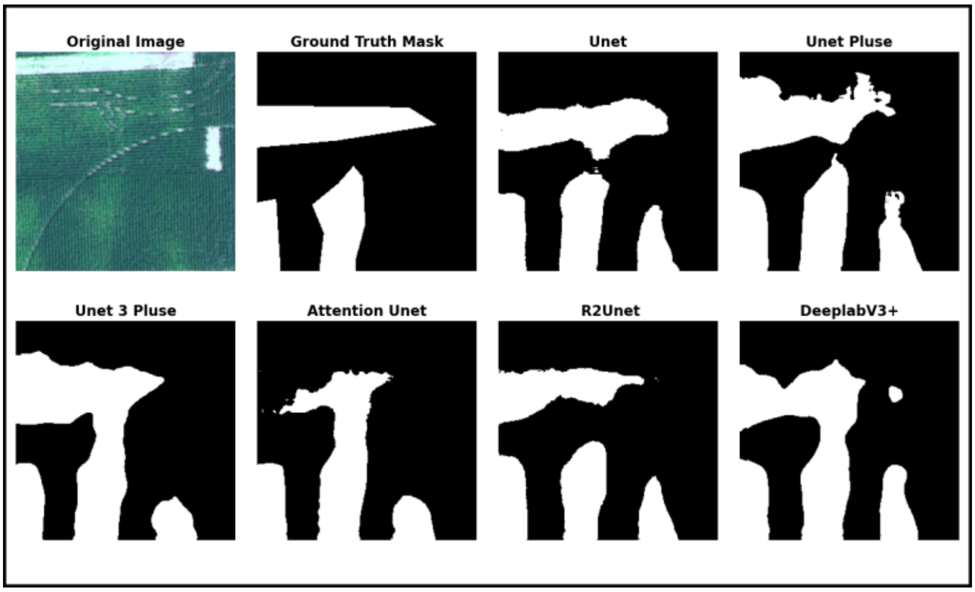


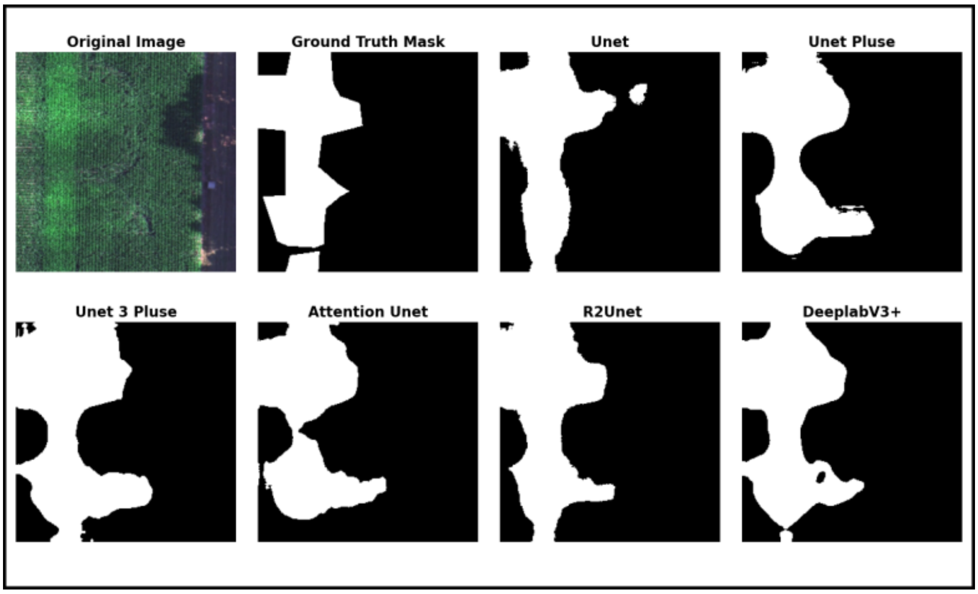




5.3.2.8 Nutrition Deficiency

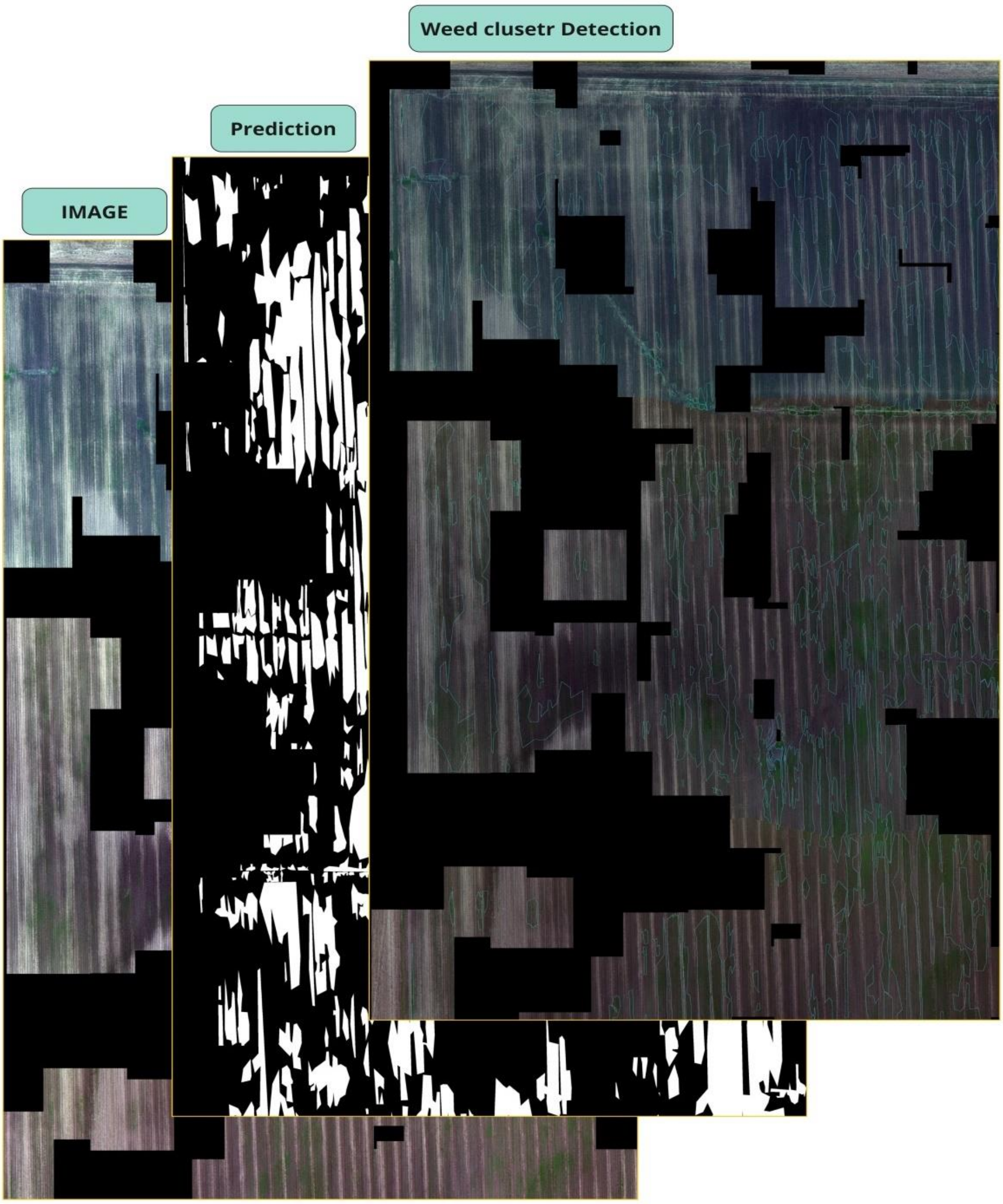




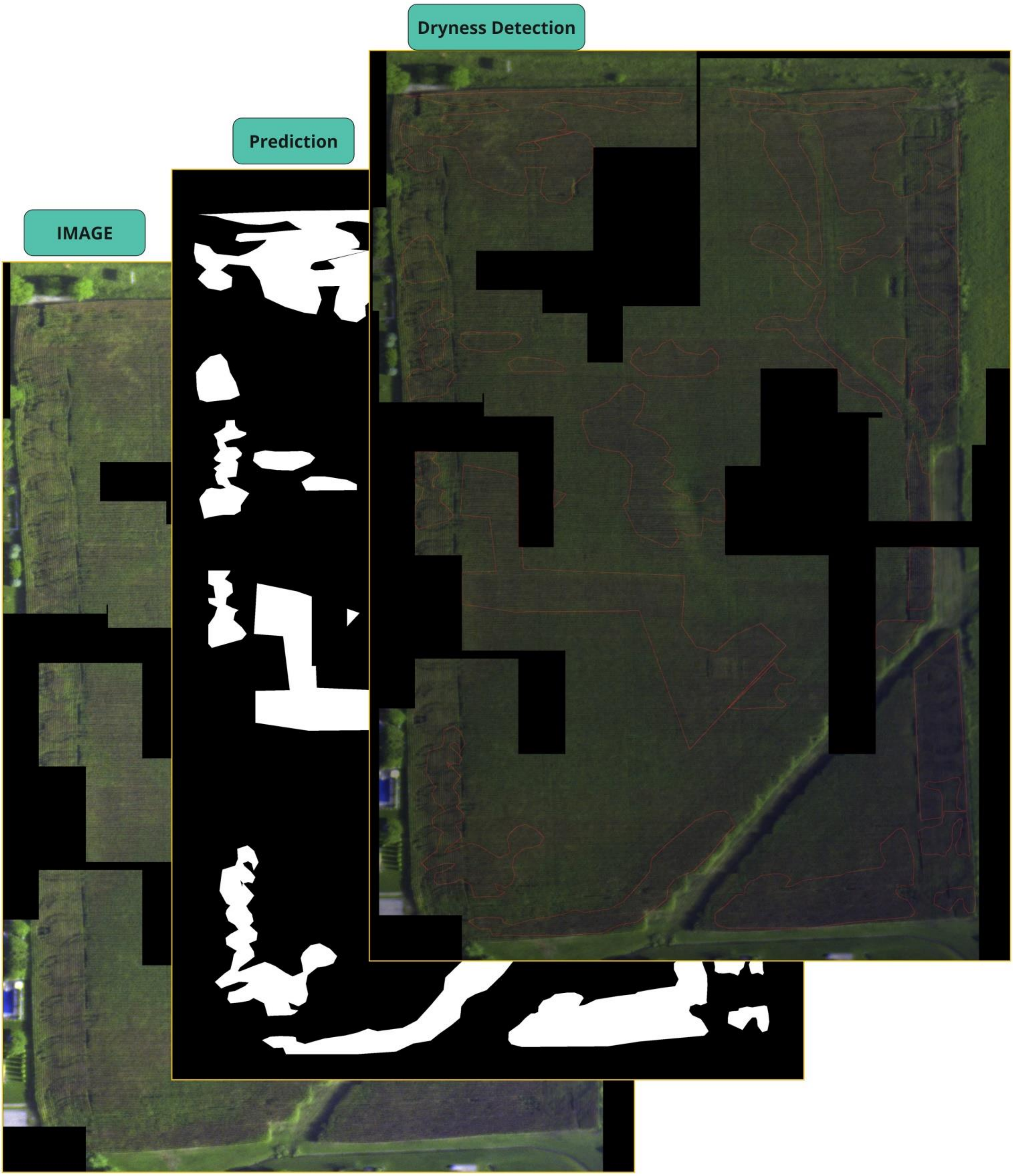


5.3.3 Prediction in Large Scale

In this project, our main focus is to empower our solution for large-scale implementation and accurately predict the affected area. By scaling up our solution, we can effectively analyze remote sensing images over extensive geographic areas and provide precise information on the impacted reign.



The farmland (QEQNXLHLY) is a 463 square kilometer area, with weed clusters as a major issue. To address this problem, the image of the farmland is divided into smaller patches (256x256 pixels) and analyzed using a suitable model to identify and classify weed clusters. The individual patch predictions are then combined to reconstruct the full-size image. The affected areas are visualized by overlaying contours or highlights, revealing a total weed-affected area of **73 square kilometers.**


Dryness Detection
Prediction
IMAGE

5.4 5. Conclusion

In conclusion, the integration of aerial imagery and deep learning techniques in agriculture offers significant potential for precise detection of anomalies in crop patterns. The utilization of advanced models such as U-Net++, Recurrent U-Net, and Recurrent Residual U-Net with Attention gate improves semantic segmentation accuracy. This enables timely intervention and optimized resource allocation, leading to improved crop health and productivity. The adoption of precision agriculture practices, supported by deep learning, contributes to sustainable and efficient crop management, addressing global challenges in food production. Continued research in this field can further enhance anomaly detection and decision support systems, benefiting farmers worldwide.

7 WATER RESOURCES ANALYSIS

ABSTRACT

Analysis of water resources and long-term trends in all regions of the earth is important for understanding the impact of climate change, particularly in monitoring drought conditions. Private households, agriculture, and other industries heavily rely on stable water supplies, which make it necessary to monitor and predict changes in water resources. In general, remote sensing technology plays a crucial role in water resource analysis, allowing us to monitor, understand, and predict changes in water resources data from different satellite missions and **LDAS**. By developing and utilizing advanced analytical tools and models, we can gain valuable insights into the state of water resources globally and support decision-making processes for sustainable development. The **SMAP** and **GRACE** missions have proven to be invaluable tools in water resources analysis and drought monitoring, providing valuable information for water resource management. By combining remote sensing data with other data sources such as satellite-based rainfall estimates, surface temperature, vegetation indices, atmospheric data, and ground measurements, comprehensive drought monitoring tools can be developed, providing crucial insights into the state of water resources, and guiding effective water resource management strategies. While limitations of remote sensing data should be considered, its importance in water resource analysis and drought monitoring cannot be overstated. With the growing scarcity of water resources, the use of remote sensing data is likely to become even more important in the coming years, providing crucial insights into the state of water resources and guiding effective water resource management strategies, including early warning systems, drought impact assessment, and decision-making support.

The drought monitor is an essential tool for managing the impacts of drought. It provides early warning, informs resource management decisions, helps with disaster preparedness, and supports policy development. With increasing frequency and severity of droughts due to climate change, the drought monitor is becoming even more critical for mitigating the impacts of drought and ensuring the sustainable use of natur**e.**

1. Introduction

Water is a vital resource for human life, supporting ecosystems and agricultural production. However, water resources are facing increasing pressure due to population growth, climate change, and unsustainable water use practices. To effectively manage water resources, it is necessary to monitor and understand how water is distributed across the planet. One important tool for this is the Soil Moisture Active Passive (**SMAP**) mission.

NASA launched the SMAP mission in 2015, designed to provide global measurements of soil moisture with high spatial accuracy. The mission uses active and passive microwave sensors to measure soil moisture, where the active sensor emits microwave pulses, and the passive sensor measures the natural microwave emissions from the Earth's surface. By combining these measurements, the SMAP mission can provide extremely accurate maps of soil moisture, which are essential for monitoring drought conditions and managing water resources.

The data provided by the SMAP mission can be used to calculate a range of drought indices, which are important for monitoring and managing water resources. One of these indices is the Standardized Precipitation Index (SPI), which measures the deviation of rainfall from the long-term average. By using SMAP data to calculate SPI, it is possible to identify areas experiencing drought conditions and track the severity and duration of these conditions over time.

The Drought Monitor is a tool that uses a variety of data sources, including SMAP data, to provide accurate and timely information about drought conditions throughout the world. The Drought Monitor is produced weekly and provides an overview of current drought conditions, as well as trends over time. The tool is used by a range of stakeholders, including farmers, water managers, and policymakers, to make decisions about water allocation and management.

The world is generally facing many problems caused by drought, such as its significant impact on agriculture, ecosystems, and human health. On the other hand, it can lead to crop failure, wildfires, and water shortages, and can have long-term effects on soil health and water availability. By providing accurate and timely information about drought conditions, the Drought Monitor enables stakeholders to take proactive measures to mitigate the effects of drought and manage water resources more effectively.

In this research paper we will use remote sensing techniques, Aws cloud techniques and data analysis techniques to analyze SMAP data to monitor soil moisture levels and detect drought conditions to help us in creating drought monitor from collect SMAP data, calculate soil moisture anomalies and generate drought maps all this step will help us to make simplified demonstration of climate change effects Through historical trends techniques, our Drought Monitor provides historical data on drought conditions, which can be used to demonstrate how droughts have become more frequent or severe over time due to climate change.

2. Literature Review

2.1 Introduction

This section demonstrates the usefulness and reliability of the SMAP L4_SM data product for various applications, including drought monitoring, hydrological modeling, and agricultural management. The studies also highlight the importance of validating the SMAP L4_SM data product using ground-based observations and other independent data sources to ensure its accuracy and reliability.

Based on a review of literature on SMAP L4 soil moisture data analysis, several methods have been identified for analyzing this dataset. These methods include:

1. Quality control and data processing: Before performing any analysis, it is important to ensure that the data is of high quality and has undergone appropriate processing. Quality control measures such as outlier detection, error correction, and data validation can be used to ensure the accuracy and reliability of the data.

2. Statistical analysis: Statistical techniques such as correlation analysis, regression analysis, and principal component analysis can be used to explore the relationships between soil moisture and other environmental variables. These techniques can also be used to identify trends and patterns in the data.

3. Spatial analysis: Spatial analysis techniques such as geostatistics and spatial interpolation can be used to generate spatially continuous maps of soil moisture. These maps can reveal the spatial patterns and variability of soil moisture across a region.

4. Machine learning and artificial intelligence: Machine learning algorithms and artificial intelligence techniques such as neural networks, support vector machines, and random forests can be used to predict soil moisture levels based on historical data and environmental variables. These models can also be used to generate future projections of soil moisture levels under different scenarios.

5. Time-series analysis: Time-series analysis techniques such as autoregressive integrated moving average (ARIMA) and wavelet analysis can be used to explore the temporal patterns and variability of soil moisture data. These techniques can also be used to identify trends and cycles in the data.

Overall, the combination of these methods allows for a comprehensive analysis of SMAP L4 soil moisture data, which can provide insights into the dynamics of soil moisture and its environmental implications.

2.2 Multi-Scale Assessment of SMAP Level 3 and Level 4 Soil Moisture Products over the Soil Moisture Network within the ShanDian River (SMN-SDR) Basin, China

The SMAP L4 products are obtained by assimilating TB data from SMAP datasets into a modified version of the GEOS-5 LDAS algorithm using a spatially distributed ensemble Kalman filter. The SMAP L1C TB and surface meteorological controlling inputs are the principal drivers of this system. The SMAP SM retrieval algorithm is affected by various parameters such as surface temperature, vegetation, and surface roughness, which can cause uncertainties in the SM estimates. To validate the SMAP SM products, complementary methodologies such as core validation measurements, sparse network, model-derived products, other satellites datasets, and field operations are being used. This study focuses on validating SMAP SM products across the ShanDian River Basin using in situ data acquired by the SMN-SDR network and evaluating the L3 and L4 SMAP SM products compared with in situ observations. The study also evaluates SMAP SM products across the SMN-SDR Basin under different vegetation types and spatial-temporal scales.[1]

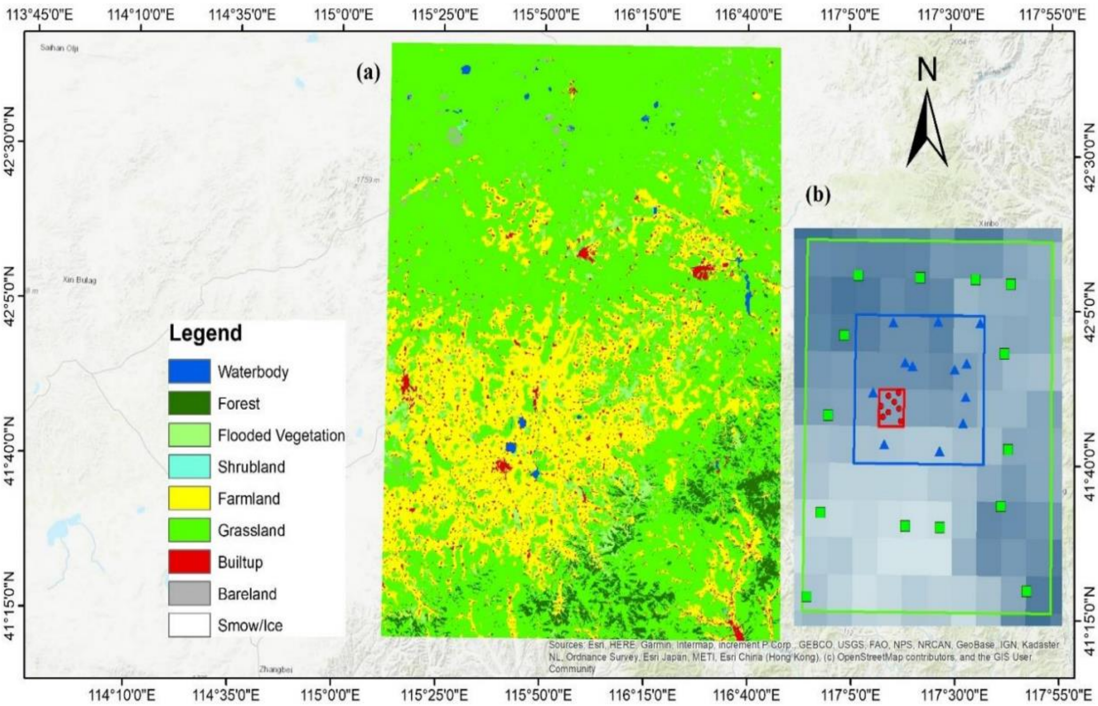
2.2.1 Materials and Methods

1. **Study Domain and Ground Observation Network and Datasets**

This study utilized in situ soil moisture data from the wireless Soil Moisture Network (SMN-SDR) within the ShanDian River Basin in the North China region. The network consisted of 34 stations with 3 sample scales and 5 measuring depths, and Decagon EM50 sensors were used to estimate SM. The data recording time period was 10 and 15 minutes, and the data sets recorded from the sensors only included raw data. The SMN-SDR also provided multi-layer soil temperature and other meteorological variables such as precipitation. The SMN-SDR in situ SM data was already calibrated and validated with ground truth SM data. In this research work, the SMN-SDR SM data at 5-cm depths from July 2018 to December 2019 was used to validate L3 SMAP_E and L4 SMAP_GAU (SSM) data, and 50 cm was used to validate Level 4 SMAP_GAU (RZSM) data.[1][2]

2. **SMAP Soil Moisture Products**

This study evaluated two SMAP soil moisture (SM) products, Level 3 enhanced SMAP radiometric SM (SPL3SMP_E) and Level 4 SMAP SSM and RZSM (SPL4SMGP), for SM estimation from January 2018 to December 2019. SMAP captures global brightness temperature (TB) with a spatial resolution of 36 km and a temporal resolution of 3 days using an L-band radar and radiometer. The SPL3SMP_E products provide daily global SSM (5 cm depth) estimates, and the SPL4SMGP product provides instantaneous SM at 3-hourly time-averaged from assimilation system for both layers, surface (0–5 cm), and root-zone (0–100 cm). The SPL3SMP_E products were already excluded from areas with adverse climatic conditions. The NSIDC website provides free access to all datasets.[1][2]



CH7 Figure 1

The ShanDian River Basin and the Network within the ShanDian River Basin (SMN-SDR), (a) land use map of the study area, (b) providing in situ SM measurements at a small scale (red circles within the red rectangles), medium scale (blue triangles within the blue rectangles), and large scale (green squares within the rectangles).

3. Statistical Analysis

This study qualitatively evaluated SMAP soil moisture (SM) products, SPL3SMP_E and SPL4SMGP, using statistical metrics such as unbiased root mean square error (ubRMSE), correlation coefficient (R), anomaly R, and mean bias. The evaluation was conducted at 2 spatial scales: core validation sites (CVSs), which provide all stations within an SMAP product grid-cell with a 9-km resolution, and a sparse network of M-scale and L-scale stations, which provide a point-scale measurement containing only 1 station within an SMAP product grid-cell (9-km). The metrics were validated using the in-situ Soil Moisture Network (SMN-SDR) topsoil layer (0-5 cm) measurements for SPL3SMP_E and SPL4SMGP SSM estimations and the arithmetic mean of SMN-SDR network observations in the up to 50 cm soil layer for SPL4SMGP RZSM. The difference in soil depths across SMAP products had minimal impact on the evaluation, as demonstrated using limited in situ measurements in the study region.[2][3]

$$\text{Bias} = E[\theta_{est}] - E[\theta_{insitu}]$$

$$\text{RMSE} = \sqrt{E[((\theta_{est}) - (\theta_{insitu}))^2]}$$

$$\text{ubRMSE} = \sqrt{E[((\theta_{est} - E[\theta_{est}]) - (\theta_{insitu} - E[\theta_{insitu}]))^2]}$$

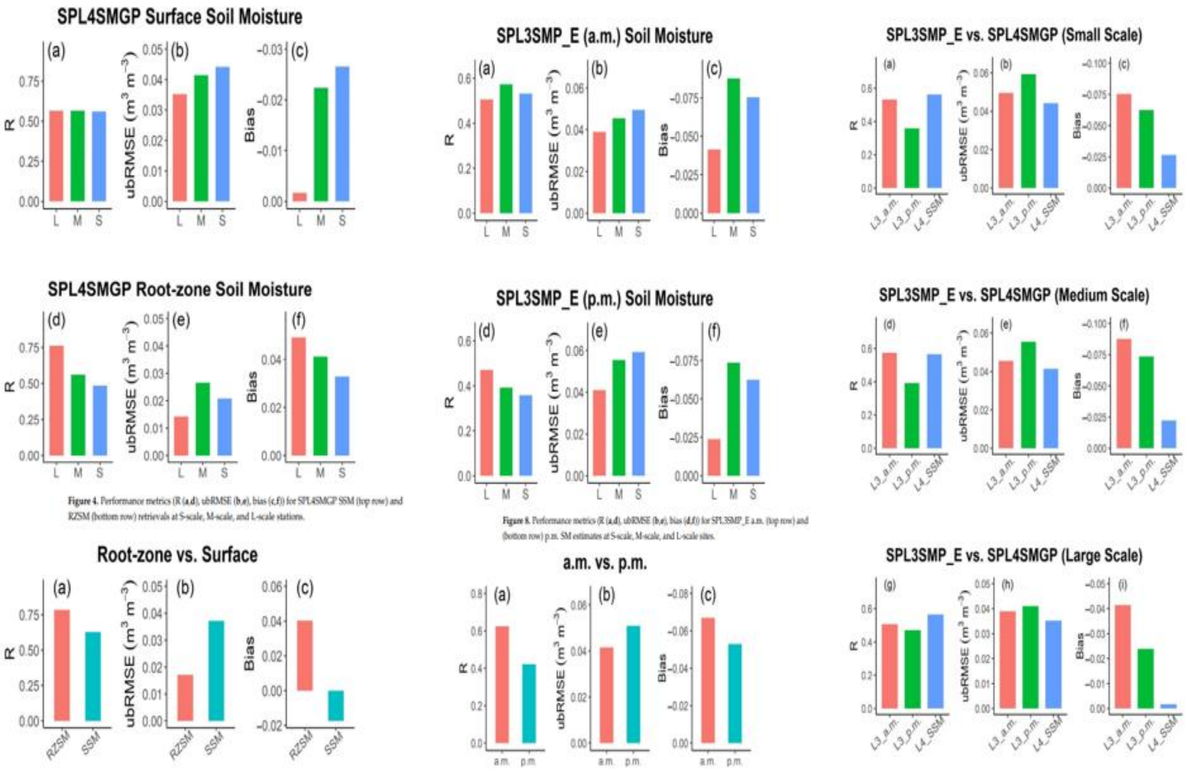
$$R = E[(\theta_{est} - E[\theta_{est}])(\theta_{insitu} - E[\theta_{insitu}])]\,(\sigma_{est}\sigma_{insitu})^{-1}$$

2.2.2 Results

1. **Evaluation of the SPL4SMGP Surface and Root-Zone Soil Moisture**
   The evaluation of the SPL4SMGP SSM and RZSM products at different spatial scales using in situ data from a total of 34 ground sites. The study found that the SPL4SMGP SSM and RZSM estimates were underestimated and overestimated, respectively, compared to the in-situ observations. The product accuracy requirement (ubRMSE value less than 0.04 m3/m3) was met by all spatial scale stations for both SPL4SMGP SSM and RZSM products. The SPL4SMGP SSM products showed a good correlation value of 0.78, while the SPL4SMGP RZSM products showed a lower correlation value of 0.63. The performance of the SPL4SMGP SM datasets over the sparse network (M and L scales) showed better results than the core validation site (S scale). Overall, the study found that the SPL4SMGP SSM and RZSM products performed well in meeting their product accuracy requirements at all spatial scales.[1][4]

2. **Evaluation of SPL3SMP_E Ascending and Descending SM**
   The accuracy of the SMAP satellite's SPL3SMP_E SM products at different spatial scales using in situ data from 34 validation sites. The study found that the SPL3SMP_E a.m. SM datasets performed well in meeting the product accuracy requirements, with good evaluation results corresponding to rainfall events. In contrast, the SPL3SMP_E p.m. SM datasets showed less accuracy, with larger errors and lower correlation values. The study also found that the satellite SM products performed better for sparse network sites than core sites, reflecting vegetation cover's adverse influence over satellite SM observations.[1][4]

3. **Comparison of SPL3SMP_E and SPL4SMGP Surface Soil Moisture**
   Compared the accuracy of the SPL3SMP_E and SPL4SMGP SSM products from the SMAP satellite at different spatial scales using in situ data from 34 validation sites. The study found that the SPL4SMGP SSM retrievals performed better than the SPL3SMP_E SM estimates, with lower ubRMSE values and higher correlation values. The study also found that the satellite SM products performed better at sparse network sites than core validation sites. Overall, the study highlights the benefits of integrating satellite-based SM and LSM using data assimilation techniques to produce optimized results.[1][4]



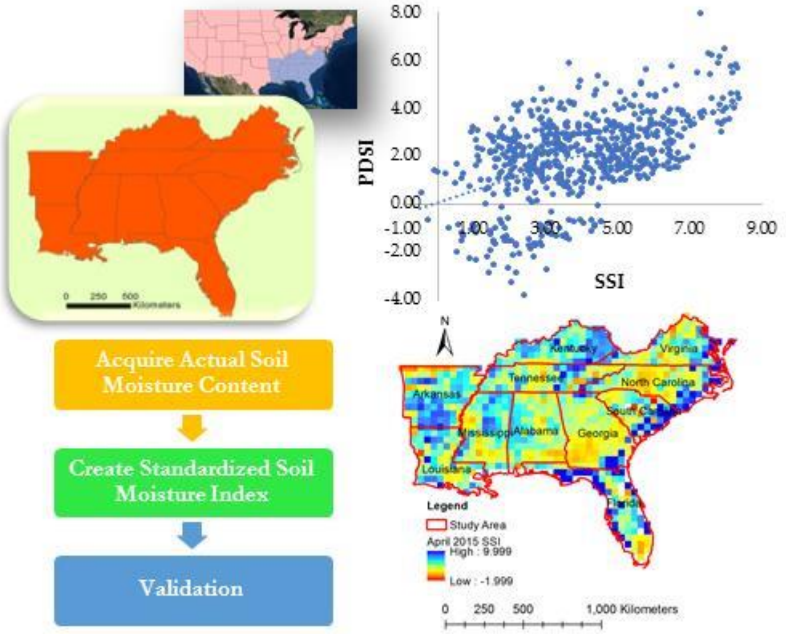
CH7 Figure2

Show the Evaluation of the SPL4SMGP Surface and Root-Zone Soil Moisture, Evaluation of SPL3SMP_E Ascending and Descending SM, Comparison of SPL3SMP_E and SPL4SMGP Surface Soil Moisture

2.3 Standardized Soil Moisture Index for Drought Monitoring Based on Soil Moisture Active Passive Observations and 36 Years of North American Land Data Assimilation System Data: A Case Study in the Southeast United States

2.3.1 Introduction

The southeastern United States is vulnerable to regional-scale droughts caused by climate variability, which affects agriculture, forestry, and rangelands. The USDA and the Southeast Regional Climate Hub (SERCH) provide science-based knowledge and tools to help farmers, ranchers, and foresters cope with climate issues in the southeast. SERCH uses a drought mitigation tool called LIGHTS, which is driven by NOAA's Climate Prediction Center's Monthly Drought Outlook, Monthly Temperature and Precipitation Outlook, and Risk of Seasonal Climate Extremes in the US related to El Niño–Southern Oscillation (ENSO). However, LIGHTS do not include soil moisture indices in the prediction model, and integrating soil moisture data can significantly enhance the reliability and accuracy of the model. Various soil moisture retrieval methods and indices have been proposed in previous research, but they have limitations in accuracy, applicability, and scalability. To address these limitations, researchers propose a new soil moisture index called the standardized soil moisture index (SSI) for drought warning. SSI incorporates data from the Soil Moisture Active Passive (SMAP) satellite and the North American Land Data Assimilation System (NLDAS) and identifies droughts as statistical outliers in the time series of soil moisture. The integration of SMAP and NLDAS data into SERCH LIGHTS is expected to improve the extent and accuracy of drought prediction models.[5]



CH7 Figure 3

2.3.2 Materials and Methods

1. **Data Acquisition**

Here we utilized three different datasets to develop a new soil moisture index for drought warning. The first dataset was obtained from the Level 3 soil moisture data from L-Band Radiometer on board the NASA satellite (SMAP). This dataset provides direct soil moisture measurement at 6 AM local solar time in the top 5-cm layer of the soil column. The second dataset was obtained from soil moisture data from the NASA North American Land Data Assimilation System (NLDAS). This dataset measures the top 10-cm soil moisture and has a time zone of Coordinated Universal Time (UTC), which has an overall six-hour time difference compared with the SMAP local solar time. The third dataset came from the Soil Climate Analysis Network (SCAN), which collects soil moisture data from probes placed at 5.08 cm depth across the United States. The USDA National Resources Conservation Service provides the SCAN dataset. The researchers used these datasets to develop a new soil moisture index called the standardized soil moisture index (SSI) for drought warning. The SSI identifies droughts as statistical outliers in the time series of soil moisture.[5]

Platform & Sensor	Parameter	Use
SMAP Passive Radiometer	Soil moisture, Level-3, 36 km resolution	Daily measurement of soil moisture
NLDAS	Soil moisture, Noah model	Historical mean and standard deviation of soil moisture
USDA SCAN	Soil moisture	Validation

CH7 Figure4 Data Description

2. Data Processing

SMAP reached its orbit in January 2015, and the data were available since 1 April 2015. Therefore, only less than two years of data have been recorded at the time of this study. The pre-processing of the SMAP data removed invalid values and outliers. The units of SMAP and NLDAS soil moisture do not match. SMAP measures the volume of water per unit volume of soil.

NLDAS measures soil moisture in units of kilogram per square meter of soil over variable thicknesses. Next equation converts the unit of NLDAS to the volume ratio that is similar to SMAP units.

$$\frac{SM_{NLDAS}\ (kg/m^2)}{W \times T}$$

This study noticed the inconsistency of the soil depth measured by the NASA North American Land Data Assimilation System (NLDAS) and the Level 3 soil moisture data from L-Band Radiometer on board the NASA satellite (SMAP). The NLDAS measures the top 10 cm of the soil, 5 cm deeper than that of SMAP. To address this inconsistency, the researchers used a linear transformation to calibrate the two datasets. The calibration coefficients between NLDAS and SMAP are listed in Table A1, which can be found in Appendix A. The researchers also used a formula that considers the original soil moisture value, density of water, and thickness of soil to calculate the standardized soil moisture index (SSI) for drought warning.[5]

3. Data Analysis

For each Julian day, there are 36 NLDAS observations from the past 36 years. We were able to calculate the mean and standard deviation of each day. The daily SSI was calculated with Equation (2):

$$SSI = \frac{x_{SMAP} - \mu_{NLDAS}}{\sigma_{NLDAS}}$$

Where x_{SMAP} is the soil moisture content from SMAP Level 3 data for a single day, μ_{NLDAS} is the mean value of soil moisture content for the corresponding day from NLDAS, and σ_{NLDAS} is the standard deviation.

4. Validation

The SMAP mission specifies the accuracy of soil moisture to be within 0.04 (4%) m^3/m^3 volumetric in low or moderately vegetated areas in the following conditions.

- Vegetation water content ≤ 5 kg/m^2
- Urban fraction ≤ 0.25
- Water fraction ≤ 0.1
- Digital Elevation Model (DEM) slope standard deviation ≤ 3 degrees

This study aimed to validate the soil moisture product obtained from the Level 3 soil moisture data from L-Band Radiometer on board the NASA satellite (SMAP) and the NASA North American Land Data Assimilation System (NLDAS) in the southeastern United States. The researchers selected seven Soil Climate Analysis Network (SCAN) stations located in agricultural lands, plains, or grasslands and representative of diverse weather conditions to compare the daily soil moisture data from SMAP and NLDAS to daily soil moisture data retrieved from the SCAN stations from 31 March 2015 to 16 July 2016. Additionally, the researchers compared the SCAN data and NLDAS data for 12 months, starting in January 2015 and ending in December 2015, to validate the soil moisture product.

Station ID	State Code	Station Name
2013	GA	Watkinsville #1
2024	MS	Goodwin Ck Pasture
2053	AL	Wtars
2039	VA	N Piedmont Arec
2005	KY	Princeton #1
2012	FL	Sellers Lake #1
2085	AR	Uapb-Earle

CH7 Figure 5 Soil Climate Analysis Network (SCAN) station used for validation.

SSI was validated by several soil moisture products, including PDSI and MODIS data. PDSI data for April 2015 were downloaded as NetCDF files in the WGS 1984 Geographic Coordinate System from the National Integrated Drought Information System on the U.S. Drought Portal. We derived a normalized difference water index (NDWI) from MODIS surface reflectance data.

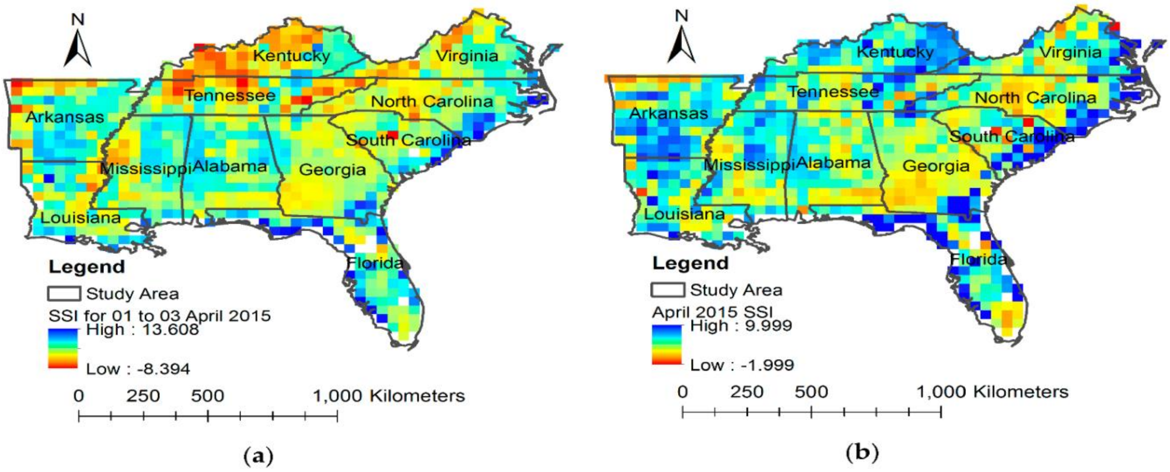
$$\mathrm{NDWI} = \frac{\mathrm{NIR} - \mathrm{SWIR}}{\mathrm{NIR} + \mathrm{SWIR}}$$

Where NIR is the near infrared reflectance and SWIR is the short-wave infrared reflectance of the MODIS data. Both PDSI and NDWI data were resampled to 36 km for the SSI validation.[5]

2.3.3 Results

1. SSI Spatial Analysis

This study used SMAP data to develop a standardized soil moisture index (SSI) for drought warning. The SSI is a z-score that measures how many standard deviations a SMAP value is from the historic mean. The SSI results were visualized using a color scale, where yellow to red colors indicate negative z-scores, and green to blue colors indicate positive z-scores.



CH7 Figure 6 Mosaic of the three consecutive standardized soil moisture index (SSI) maps from 1 to 3 April 2015. Areas in yellow to red represent areas that are experiencing very dry conditions, indicating drought. (**b**) SSI map for the whole month of April 2015.

2. Validation with PDSI and NDWI

PDSI is a standardized index that spans −10 (dry) to +10 (wet) [34].

Areas in yellow and red represent areas that are experiencing dry conditions; (b) Normalized difference water index (NDWI) calculated for 01 to 03 April 2015. Likewise, areas in yellow and red represent areas that are experiencing low vegetation water content and therefore a dry condition.[5]

3. Methodology

In this section, we provide a comprehensive overview of the methods we used to analyze water resources, especially soil moisture data analysis, and to create a drought monitoring system that can be used to monitor and manage water resources during periods of drought. Our methodology includes a range of scientific and technical approaches, including data processing, quality control, modeling, validation, spatial analysis, and data visualization. These techniques are used to provide valuable information for drought monitoring, flood prediction, and water resource management. Our methodology can also be adapted to different regions and applications, providing a flexible and robust tool for soil moisture science and drought management.

3.1 Dataset

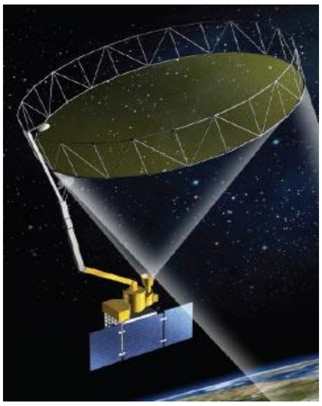
Our primary objective is to detect and analysis water resources, and to achieve this, we emphasize the use of soil moisture data. We have identified several techniques to acquire and extract this data, including:

1. **Ground-based sensors**: These sensors are installed in the ground to directly measure the moisture content. They can be placed at different depths and locations to provide a spatially distributed measurement of soil moisture.
2. **Satellite data:** Satellite-based sensors, like the SMAP mission, offer global coverage of soil moisture levels. They use microwave radiation to estimate soil moisture levels and provide data at regular intervals.
3. **Aircraft-based sensors:** These sensors are installed on aircraft and can measure soil moisture levels from the air. They provide high-resolution data over specific areas and can be used for research or monitoring purposes.
4. **Portable soil moisture meters:** These meters can measure soil moisture levels at specific locations and are easy to use. They provide quick measurements of soil moisture levels.
5. **Weather stations:** Some weather stations are equipped with soil moisture sensors that measure soil moisture levels in addition to other weather variables, such as temperature, humidity, and precipitation.
6. **Soil maps:** These maps provide estimates of soil moisture levels based on soil type and other soil properties. They can be used to estimate soil moisture levels for large areas.

The choice of method will depend on factors such as the level of accuracy required, the spatial and temporal resolution needed, and the resources available for data acquisition and processing.

To meet our objective of providing global coverage of soil moisture data, we used satellite data and there are two choices to get soil moisture data.[6]

5.1.1 Availability of soil moisture data

	SMAP	JAXA
Specifications	1. Parameter(s): Root zone soil moisture surface soil moisture 2. Platform(s): GEOS-5, SMAP 3. Sensor(s): not application, SMAP L-BAND RADIOMETER 4. Data Format(s): HDF5 5. Temporal Coverage:31 March 2015 to present. 6. Temporal Resolution:3 hour 7. Spatial Resolution:9 km x 9 km 8. Spatial Coverage: N:85.044, S: -85.044, E:180, W: -180	1. Parameter(s): Brightness Temperature Soil Moisture / Water content 2. Platform(s): GCOM-W1 3. Sensor(s): AMSR2 4. Data Format(s): HDF-EOS5 5. Temporal Coverage:2 July 2012 to present. 6. Temporal Resolution:50 minute 7. Spatial Resolution:25 km x 25km 8. Spatial Coverage: N:89.24, S: -89.24, E:180, W: -180
Advantages	1.High accuracy 2.High spatial resolution	1.High temporal resolution 2.Integration with other data
Disadvantages	1.Limited spatial resolution 2.Limited data availability	1.Lower accuracy 2.Lower temporal resolution
	 CH7 Figure 7: SMAP Satellite.	 CH7 Figure 8: JAXA Satellite

CH7 Table 1

we utilized satellite data from the SMAP mission. This mission was chosen due to its high accuracy and availability for data acquisition and processing and for availability of data with binary format HDF5 (Science Values) that help us in our research and Application, which align with our motivation.[9][8]

3.1.2 HDF5 (Hierarchical Data Format version 5)

is a file format for storing and organizing large and complex datasets. It is a binary format that is designed to be platform-independent, meaning that HDF5 files can be read and written on different operating systems and with different programming languages.

HDF5 files consist of two main components: datasets and attributes. A dataset is a collection of data elements organized in a multidimensional array, while an attribute is a metadata object that provides additional information about the dataset. HDF5 files can also contain groups, which are hierarchical structures that can contain datasets, attributes, and other groups.

HDF5 is commonly used in scientific and engineering applications for storing and sharing large datasets, such as satellite imagery, climate data, and simulation results. It provides a flexible and efficient way to store and access complex data, as well as a rich set of tools and libraries for working with HDF5 files in different programming languages, including C, C++, Python, and MATLAB.[7][9]

3.2 SMAP Mission Satellite

3.2.1 SMAP mission overview

The Soil Moisture Active Passive (SMAP) mission is an orbiting observatory that measures the amount of water in the surface soil everywhere on Earth.

It was launched in January 2015 and started operation in April 2015. The SMAP radiometer has been operating flawlessly. The radar instrument, ceasing operation in early 2015 due to failure of radar power supply, collected close to 3 months of science data. The prime mission phase of three years was completed in 2018, and since then SMAP has been in extended operation phase.

SMAP is designed to measure soil moisture, every 2-3 days. This permits changes, around the world, to be observed over time scales ranging from major storms to repeated measurements of changes over the seasons.

Everywhere on Earth not covered with water or not frozen, SMAP measures how much water is in the top layer of soil. It also distinguishes between ground that is frozen or thawed. Where the ground is not frozen, SMAP measures the amount of water found between the minerals, rocky material, and organic particles found in soil everywhere in the world (SMAP measures liquid water in the top layer of ground but is not able to measure the ice.)

SMAP will produce global maps of soil moisture. Scientists will use these to help improve our understanding of how water, energy, and carbon fluxes (in its various forms) maintain our climate and environment. The water cycle involves more than the obvious processes cycling through the steps of evaporation from the oceans and land to condensation forming clouds that then drop rain or snow on the ground (precipitation), followed by the water flowing across the land before returning to the sea. For example, plants absorb water from the soil to grow, but they also “transpire” some of it straight back into the air.

Frequent and reliable soil moisture measurements from SMAP will help improve the predictive capability of weather and climate models.[6][9]

3.2.2 SMAP mission Specifications [9]

a) Polar Orbit: Altitude_ 685 km
b) Spatial Coverage: Global
c) Launched Jan 31, 2015
d) Temporal Coverage: April 2015 – present
e) Sensors: Microwave Radiometer
Microwave Radar (not currently available)

Launch vehicle	Delta II 7320-10C
Launch site	Vandenberg Air Force Base, California
Launch date/time	Jan. 31, 2015
Orbit type	Circular polar orbit, 8-day repeating ground track
Orbit altitude	685 km (426 miles)
Orbit period	98.5 minutes
Spacecraft type	3-axis stabilized
Spacecraft mass	944 kg
Spacecraft downlink	Via NASA Near-Earth Network Stations, supplemented by Tracking and Data Relay Satellite (TDRS) support
Antenna size	6 meters (20 feet); offset-fed mesh reflector
Antenna mass	65.5 kg (144 lbs)
Radar mass	50.8 kg (112 lbs)
Radar frequency	1.26 GHz; non-imaging (unfocused) synthetic aperture
Radar transmit power	550 Watts (peak)
Radar real-aperture footprint	29 km x 35 km (18 miles x 22 miles)
Radar ground resolution	1-3 km (0.62-1.86 miles) after unfocused synthetic aperture processing over outer 70% of swath
Radiometer mass	30 kg (electronics only)
Radiometer frequency	1.41 GHz
Radiometer footprint	39 km x 47 km
Radiometer ground resolution	30 km (18.6 miles)
Mapping scheme	Antenna spins at 14.6 revolutions per minute, sweeping out overlapping loops 1000 km (621 miles) in diameter. Overlapping loops are combined to make wide ribbons which overlap each other. Mapping Earth's surface takes 2 days near the poles and 3 days near the equator.

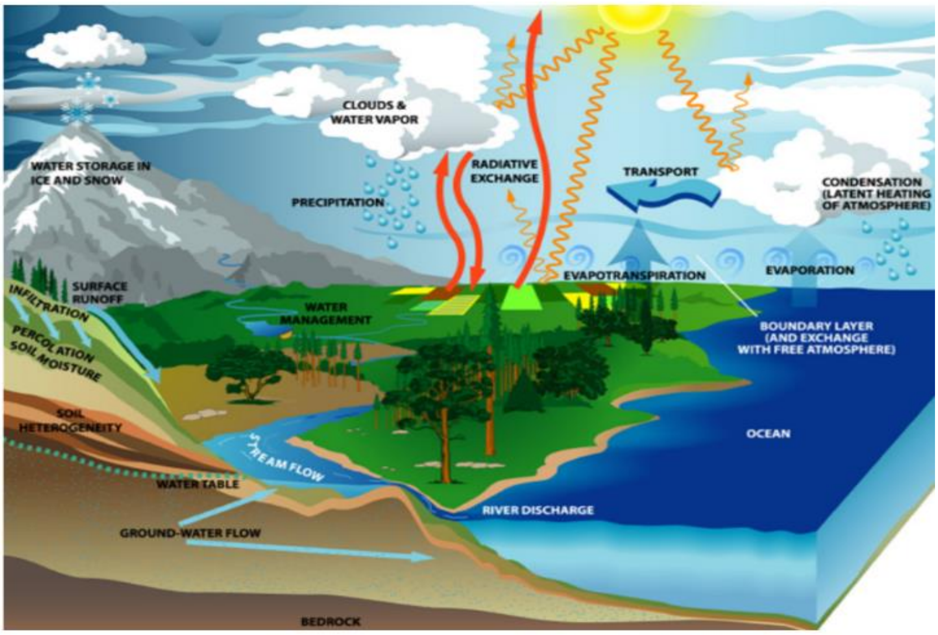
CH7 Table 2

3.2.3 SMAP mission objective

SMAP will provide a capability for global mapping of soil moisture and freeze/thaw state with unprecedented accuracy, resolution, and coverage. SMAP science objectives are to acquire space-based hydrosphere state measurements over a three-year period to:

- Understand processes that link terrestrial water, energy, and carbon cycles.
- Estimate global water and energy fluxes at the land surface.
- Quantify net carbon flux in boreal landscapes.
- Enhance weather and climate forecast skill.

- Develop improved flood prediction and drought monitoring capabilities.[6]



The water cycle is illustrated here. Water is constantly moving on Earth. The water cycle consists of all the processes involved in the transfer and storage of water in Earth's atmosphere, on its surface, underground, and by organisms living on our planet.

CH7 Figure 9

3.3 Soil Moisture Data Product

The Soil Moisture Active Passive (SMAP) L4 data products are the highest level of data products produced by the SMAP mission. These products are derived from the SMAP Level 1C and Level 2 data and provide estimates of soil moisture and vegetation water content at a spatial resolution of approximately 9 km.

There are two types of SMAP L4 data products: the Soil Moisture Analysis Ready Data (SM-ARD) and the Soil Moisture Active Passive Enhanced L4 (SMAPL4_E).

The SM-ARD product is a daily global product that provides estimates of soil moisture at a depth of 5 cm and vegetation water content. These estimates are provided in two different formats: (1) as retrievals based on a physical model that combines SMAP L1C and L2 data with ancillary data and (2) as analyses that are based on a land surface model that assimilates the retrievals along with other data sources. The SM-ARD product is available from April 1, 2015, to the present.

The SMAPL4_E product is a research product that provides estimates of soil moisture at a depth of 5 cm, surface soil moisture, and vegetation water content at a spatial resolution of approximately 3 km. This product is generated using a data assimilation technique that combines SMAP L1C and L2 data with other data sources such as precipitation and temperature data. The SMAPL4_E product is available from April 1, 2015, to the present, but it is not yet an operational product.

Overall, these SMAP L4 data products provide valuable information on soil moisture and vegetation water content for a range of applications, including hydrologic modeling, drought monitoring, and agricultural management.[9]

Product	Description	Gridding (Resolution)	Latency	
L1A_TB	Radiometer Data in Time-Order	-	12 hrs	Instrument Data
L1A_S0	Radar Data in Time-Order	-	12 hrs	
L1B_TB	Radiometer T_B in Time-Order	(36x47 km)	12 hrs	
L1B_S0_LoRes	Low Resolution Radar σ_o in Time-Order	(5x30 km)	12 hrs	
L1C_S0_HiRes	High Resolution Radar σ_o in Half-Orbits	1 km (1-3 km)	12 hrs	
L1C_TB	Radiometer T_B in Half-Orbits	36 km	12 hrs	
L2_SM_A	Soil Moisture (Radar)	3 km	24 hrs	Science Data (Half-Orbit)
L2_SM_P	Soil Moisture (Radiometer)	36 km	24 hrs	
L2_SM_AP	Soil Moisture (Radar + Radiometer)	9 km	24 hrs	
L3_FT_A	Freeze/Thaw State (Radar)	3 km	50 hrs	Science Data (Daily Composite)
L3_SM_A	Soil Moisture (Radar)	3 km	50 hrs	
L3_SM_P	Soil Moisture (Radiometer)	36 km	50 hrs	
L3_SM_AP	Soil Moisture (Radar + Radiometer)	9 km	50 hrs	
L4_SM	Soil Moisture (Surface and Root Zone)	9 km	7 days	Science Value-Added
L4_C	Carbon Net Ecosystem Exchange (NEE)	9 km	14 days	

CH7 Table 3: SMAP product

Table4 specifies several important characteristics of the SMAP baseline data products. Among these are:

- the product short name – these names provide a shorthand method to differentiate among the products.
- a very brief product description – additional SMAP project documents provide complete specifications of the format and the content of each data product; the SMAP DAACs have made these documents available to the user community.
- product spatial resolution or grid posting – the resolution of some Level 1 products is based on the size of the Instantaneous Field of View (IFOV) of the radar or radiometer instrument, while the resolution of higher-level products is based on the selected grid spacing.
- product latency to the user community – latency measures the time between the acquisition of the first element in the data product and the time the product is available for use at one of the SMAP Data Centers. Although the latencies listed in Table 2 are being used by the project to construct the data processing system, the SMAP project will do its best to deliver products sooner, whenever possible [9]

3.3.1 SMAP L4_SM

Soil Moisture [Surface and Root Zone]

3.3.1.1 Main objective of SMAP L4_SM product

1. to provide estimates of root zone soil moisture (defined here nominally as soil moisture in the top 1 m of the soil column) based on SMAP observations.
2. to provide a global surface and root zone soil moisture product that is spatially and temporally complete.

These objectives address two limitations of the SMAP Level 2 soil moisture products, which provide soil moisture estimates only for the surface layer (~top 5 cm of the soil) and only at times and locations where soil moisture can be observed by SMAP sensors (subject to orbit and land surface characteristics)

3.3.1.2 Limitation of SMAP L4 data products for soil moisture analysis

While SMAP L4 data products provide valuable information on soil moisture and vegetation water content, there are several limitations that must be considered when using these data products for soil moisture analysis. Some of these limitations are:

1. Spatial resolution: The spatial resolution of SMAP L4 data products is approximately 9 km for the SM-ARD product and 3 km for the SMAPL4_E product. This may not be sufficient for some applications that require higher spatial resolution data, such as agricultural management or urban hydrology.

2. Temporal resolution: SMAP L4 data products are available daily, but they may not capture the short-term variability of soil moisture that occurs during precipitation events or irrigation events.

3. Vegetation effects: SMAP L4 data products provide estimates of both soil moisture and vegetation water content, but the accuracy of the vegetation water content estimates may be affected by the presence of dense vegetation cover.

4. Soil texture effects: SMAP L4 data products provide estimates of soil moisture at a depth of 5 cm, but the accuracy of these estimates may be affected by the soil texture. Soils with high clay content, for example, may have different soil moisture dynamics compared to soils with high sand content.

5. Data gaps: SMAP L4 data products may have gaps due to cloud cover, data processing errors, or other issues. These gaps can limit the usefulness of the data for some applications.

It is important to consider these limitations when using SMAP L4 data products for soil moisture analysis and to evaluate the suitability of these data products for specific applications. Additionally, there may be opportunities to address some of these limitations using complementary data sources or through the development of new analysis methods.[10][11]

3.3.1.3 Product Characteristics [10][11]

The baseline L4_SM product is derived from the downscaled (9 km) brightness temperatures provided with the Level 2 Radar and Radiometer Soil Moisture (L2_SM_AP) product.

1. Geophysical parameters

a) Surface soil moisture (0-5 cm vertical average).
b) Subsurface (or "root zone") soil moisture (0-100 cm vertical average).
c) Additional research products (not validated), including surface meteorological forcing variables, soil temperature, evaporative fraction, net radiation, etc. and error estimates for select output fields that are produced internally by the L4_SM algorithm.

2. Spatial resolution, posting, and coverage

The geophysical parameters of L4_SM will be obtained and posted on the SMAP Earth-fixed global grid with a resolution of 9 km, in accordance with the EASE (version 2) grid used for other SMAP products, as described in Reichle et al. (2014b).

3. Temporal resolution and sampling

a) the land model computational time step (7.5 min).
b) the EnKF analysis update time step.
c) the reporting (or output) time step for the instantaneous and time average geophysical fields that are stored in the L4_SM data product.

The EnKF analysis update step will assimilate the available SMAP observations at the nearest 3-hourly analysis time, which includes 0z,3z, and so on up to 21z. Geophysical parameters will be provided as 3-hourly averages between these update times, according to Reichel et al. (2014b).

4. Latency

The L4_SM product must be produced within 7 days of satellite data acquisition after the 3-month In-Orbit Checkout period of the SMAP observatory. Output will be delivered once daily with a mean latency of approximately 3 days based on the availability of daily-average precipitation observations. A beta-version L4_SM product must be delivered 6 months after IOC, and the validated product will be delivered after the 12-month Calibration/Validation phase, which covers the first twelve months after IOC.

5. Error estimate

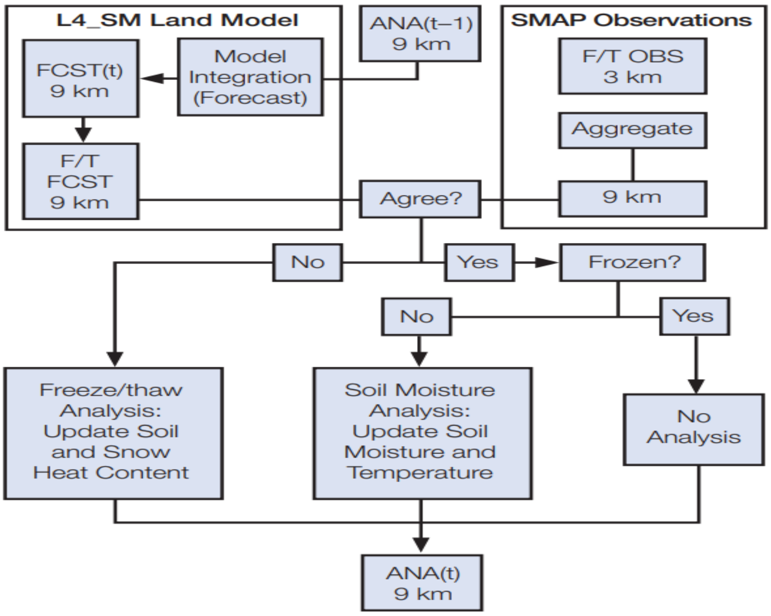
The data assimilation system dynamically determines error estimates of the assimilation product by weighing the relative errors of assimilated SMAP brightness temperature observations and the corresponding land model forecast. The accuracy of the error estimates depends on the input error parameters and needs to be validated. Instantaneous error fields will be provided for selecting variables as part of the Analysis Update Data File Collection, derived from the ensemble standard deviation of the analyzed fields, and varying in space and time.

3.3.2 L4_SM Algorithm

The L4_SM algorithm is based on two main components adapted from the Goddard Earth Observing Model System, Version 5 (GEOS-5):

a) the GEOS-5 Catchment land surface model, which provides a numerical representation of the water and energy transport processes at the land-atmosphere interface, including a land surface microwave radiative transfer model.
b) the GEOS-5 ensemble-based land data assimilation system.

The latter merges SMAP observations with estimates from the land model, driven by observation-based surface meteorological forcing data that includes a soil moisture analysis based on the ensemble Kalman filter and a rule-based freeze/thaw analysis. Assimilation of downscaled (9-km) brightness temperatures (L2_SM_AP) will occur when and where available, supplemented with 36-km brightness temperature observations (L1C_TB; ascending and descending passes) when downscaled data are unavailable. Additionally, 3-km freeze/thaw observations (L3_FT_A) will also be assimilated.



CH7 Figure 10: SM_L4 Algorithm Overview

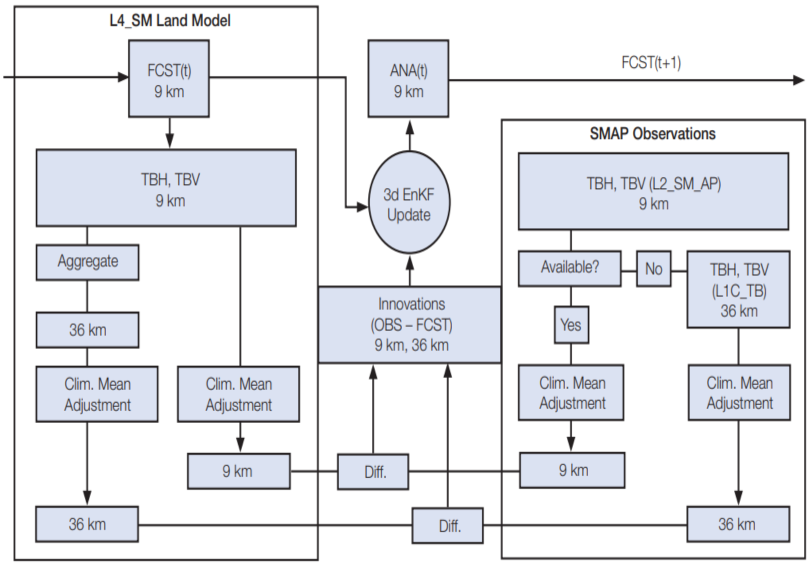
In Figure 10, an outline of one forecast and analysis cycle is presented. The algorithm starts with a Catchment model ensemble forecast, which is initialized with the analysis at time t-1 and valid at time t (labeled FCST(t) in Figure 4). For each 9 km model grid cell, the forecast freeze/thaw (F/T) state is first compared to the corresponding SMAP freeze/thaw observations, which are aggregated to the resolution of the model forecast.

If there is a discrepancy between the Catchment model forecast and the SMAP observations, the model's states in the grid cell in question are corrected towards the observations in a freeze/thaw analysis. If the forecast and observed freeze/thaw states agree and indicate non-frozen conditions, the grid cell in question is included in a distributed soil moisture analysis.

When the model indicates non-frozen conditions, but freeze/thaw observations are not available, the grid cell is also included in the soil moisture analysis. On the other hand, if the analysis step is not required for the grid cell in question, it is skipped. Once the analysis has been completed for all grid cells, the algorithm proceeds with a model forecast to time t+1, and so on. [10][11]

3.3.3 Soil Moisture Analysis

the L4_SM product is produced within 7 days of satellite data acquisition after the 3-month In-Orbit Checkout period of the SMAP observatory. The product is delivered on a 9-km Equal-Area Scalable Earth Grid (EASE-Grid 2.0) and includes uncertainty estimates of the soil moisture estimates. Quality control flags are also included in the product to indicate the validity of the data in each grid cell.



CH7 Figure 11: Soil Moisture Analysis

This flowchart explains the process of analysis of soil moisture.

a) The L4_SM product is derived using the Catchment model, which assimilates brightness temperature observations from the SMAP instrument to estimate soil moisture. The process starts with obtaining the brightness temperature observations from SMAP and downscales them to a 9-km grid cell resolution.

If downscaled data are unavailable, 36-km brightness temperature observations are used instead. The observed and modeled brightness temperatures are then adjusted by subtracting their respective seasonally varying, climatological mean values before computing the innovations vector.

b) The data assimilation system weighs the relative errors of the assimilated SMAP brightness temperature observations and the corresponding land model forecast to dynamically determine error estimates of the assimilation product. Estimates of the error of the assimilation product are then provided for select variables as part of the Analysis Update Data File Collection.

c) The state vector for the soil moisture analysis consists of seven Catchment model prognostic variables at each 9-km grid cell. The increments are computed in the units of the Catchment model prognostic variables using three-dimensional updates, and at least 24 ensemble members will be used to reduce sampling errors. [10][11]

$$x^- = \begin{bmatrix} x_1^- \\ x_2^- \\ \cdots \\ x_{N9}^- \end{bmatrix}, \quad \text{where} \quad x_j^- = \begin{bmatrix} \mathrm{SRFEXC\_FCST}_j \\ \mathrm{RZEXC\_FCST}_j \\ \mathrm{CATDEF\_FCST}_j \\ \mathrm{TC1\_FCST}_j \\ \mathrm{TC2\_FCST}_j \\ \mathrm{TC4\_FCST}_j \\ \mathrm{GHT1\_FCST}_j \end{bmatrix},$$

N9 is the number of 9-km grid cells included in the soil moisture analysis, and j=1...N9. For clarity, the subscripts for time and ensemble member are omitted.

3.4 SMAP L4 Data Access

There are different ways for access SM_L4

1. National Snow and Ice Data Center (NSIDC) DAAC: The NSIDC DAAC is the primary data archive and distribution center for SMAP data products. You can access and download SMAP L4 data products from the NSIDC DAAC website (**https://nsidc.org/data/smap**). The website provides various tools and services for searching, browsing, and downloading SMAP data products.

2. NASA Earthdata Search: The NASA Earthdata Search tool (https://search.earthdata.nasa.gov) is another way to access and download SMAP L4 data products. The tool allows you to search and filter SMAP data products based on various parameters, such as date, location, and product type. You can download the data products directly from the search results or add them to a cart for bulk download.

3. Application Programming Interfaces (APIs): You can also access SMAP L4 data products programmatically using APIs provided by the NSIDC DAAC or other third-party services. For example, the NSIDC DAAC provides a RESTful API for accessing SMAP data products (https://nsidc.org/data/smap/data-access). You can use this API to automate data retrieval and processing tasks using scripting languages such as Python or R.

4. Command-line tools: The NSIDC DAAC also provides command-line tools for accessing and downloading SMAP data products, such as the NSIDC Data Access and Management System (NDAMS) and the Simple Subset Wizard (SSW) tool. These tools can be useful for automating data retrieval and processing tasks on remote servers or cloud computing environments.[12]

3.5 DROUGHT MONITOR

3.5.1 Drought Problem Definition

A drought is a prolonged period of abnormally dry weather or precipitation deficiency, resulting in a water shortage that can adversely affect human activities, agriculture, and the environment. Droughts can be caused by various factors, including climate variability and change, natural weather patterns, and human activities such as overuse of water resources.[13][14]

3.5.2 Impact of Drought

1. Crop and pasture losses: Drought can reduce soil moisture, which can lead to reduced crop yields and pasture productivity. Plants may also become more vulnerable to pests and diseases during drought conditions.

2. Livestock losses: Drought can lead to a shortage of forage and water for livestock, which can result in reduced weight gain, increased mortality, and decreased reproductive performance.

3. Reduced water supplies: Drought can lead to decreased streamflow, reduced water levels in lakes and reservoirs, and decreased groundwater recharge, which can reduce water supplies for irrigation, industry, and domestic use.

4. Wildfires: Drought can increase the risk of wildfires, which can have significant impacts on ecosystems and communities.

5.Water quality: Drought can lead to reduced water quality, as low streamflow can concentrate pollutants and increase the risk of algal blooms.

6.Economic impacts: Drought can have significant economic impacts, including reduced agricultural productivity, increased food prices, and reduced revenues for industries that rely on water, such as hydroelectric power and recreation.

7.Social impacts: Drought can have significant social impacts, including increased food insecurity, reduced access to water for drinking and sanitation, and increased conflicts over water resources.

Mitigating the impacts of drought requires a combination of water conservation measures, drought-resistant crops and livestock, and effective drought monitoring and management strategies.

By implementing these strategies, communities and water managers can help to reduce the impacts of drought and ensure that water resources are managed in a sustainable and resilient manner.[13][14]

3.5.3 Drought Monitor Definition

The Drought Monitor is a tool used in the United States to track and monitor drought conditions across the country. It is produced by a partnership of government agencies, including the National Oceanic and Atmospheric Administration (NOAA), the U.S. Department of Agriculture (USDA), and the National Drought Mitigation Center (NDMC) at the University of Nebraska-Lincoln.

The Drought Monitor uses a combination of indicators, including precipitation, soil moisture, streamflow, and vegetation health, to assess the severity of drought conditions in different regions of the country. The data is collected and analyzed on a weekly basis, and the results are presented in a map that shows the extent and severity of drought conditions across the United States. The map is color-coded to indicate different stages of drought, ranging from abnormally dry to exceptional drought, with corresponding impacts on agriculture, water supplies, and other sectors. The Drought Monitor is used by policymakers, water managers, and the public to make informed decisions about water use, drought preparedness, and response to drought emergencies.[15]

3.5.4 Category Scale of Drought Monitor

The Drought Monitor uses a five-category scale to classify drought severity, each with corresponding impacts on agriculture, water resources, and other sectors. The categories, from least severe to most severe, are:

1. Abnormally Dry (D0): This indicates that an area is experiencing dryness, but not yet at drought levels. The impacts are minimal, but there may be some concerns for agriculture and wildfire risk.
2. Moderate Drought (D1): This indicates that an area is experiencing some degree of drought, with some damage to crops and pastures, and an increase in the risk of wildfires.
3. Severe Drought (D2): This indicates that an area is experiencing significant drought conditions, with major damage to crops, pastures, and water supplies. Water restrictions may be necessary, and there is an increased risk of wildfires.
4. Extreme Drought (D3): This indicates that an area is experiencing severe and widespread drought conditions, with major crop and pasture losses, widespread water shortages, and an increased risk of water conflicts and wildfires.
5. Exceptional Drought (D4): This indicates that an area is experiencing exceptional and widespread drought conditions, with complete crop and pasture losses, water shortages in reservoirs, streams, and wells, and widespread water emergencies.

This category indicates a water supply crisis, with impacts on agriculture, industry, and communities.[16][17]

3.6 Worldwide satellite data-based Drought Monitor

Our goal is to develop a global Drought Monitor that can aid in predicting drought indices and promptly identifying solutions.[16][17]

It’s included:

- Daily update
- Comparable Measurement worldwide
- Comparable resolution worldwide
- simplified demonstration of climate change effects

3.6.1 SM_L4 Data for Creating Drought Monitor

Creating a drought monitor from soil moisture data involves several steps. Here are the main steps:

1. Obtain soil moisture data: You can obtain soil moisture data from various sources, such as satellite data, ground-based sensors, or model output. The data should cover the desired region and time period.

2. Calculate the soil moisture anomalies: Calculate the deviation of the current soil moisture from the long-term average soil moisture in the region of interest. This can be done by subtracting the long-term average soil moisture from the current soil moisture data.

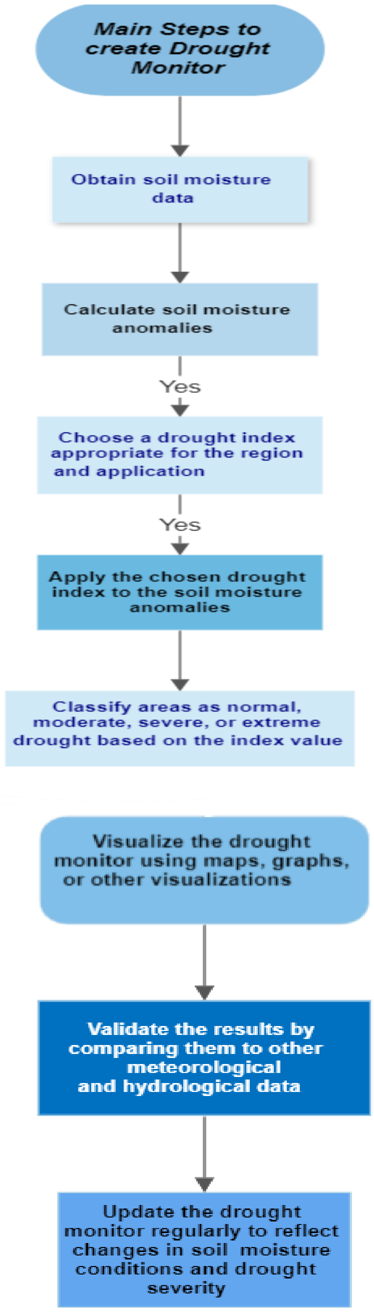
3. Choose a drought index: Choose a drought index that is appropriate for the specific region and application. Some common drought indices include the Standardized Precipitation Index (SPI), the Palmer Drought Severity Index (PDSI), the Standardized Soil Moisture Index (SSI), and the Soil Moisture Anomaly (SMA).

4. Apply the drought index to the soil moisture anomalies: Apply the chosen drought index to the soil moisture anomalies to create a drought monitor. The drought monitor will classify areas as being in normal, moderate, severe, or extreme drought based on the index values.

5. Visualize the drought monitor: Visualize the drought monitor using maps, graphs, or other visualizations to make it easy to interpret the data. The visualization should highlight the areas that are experiencing drought and the severity of the drought.

6. Validate the results: Validate the results of the drought monitor by comparing them to other meteorological and hydrological data, such as precipitation, streamflow, and vegetation health. This will help to ensure the accuracy and reliability of the drought monitor.

7. Update the drought monitor: Update the drought monitor regularly to reflect changes in the soil moisture conditions and drought severity in the region. This will help to provide timely and accurate information to decision-makers and stakeholders.



CH7 Figure 12 Drought Monitor Flowchart

It's important to note that creating a drought monitor from soil moisture data requires expertise in both data analysis and drought monitoring. It's also important to consider other factors that can contribute to drought, such as temperature, precipitation, and vegetation health, when interpreting the results of a drought monitor.[18][19]

3.6.2 SM_L4 Soil Moisture Data Analysis [20][21]

- First, using these tools to create request Algorithm for Sm_L4:
1. python program language
2. AWS Sage Maker tutorial cloud
3. Anaconda software

- Main arguments in algorithms:

1.Start date
2.End date
3.Bounding box
4.Short name for data
5.Version
6.CMR_URL
7.URS_URL

CH7 Table 4: SMAP Data argument

- Downloading Data with Binary Format HDF5 files then reading data we need from files

Soil Moisture Data from HDF 5
Analysis_Data
EASE2_global_projection
Forecast_Data
Metadata
Observations_Data
Cell_column
Cell_lat
Cell_lon
Cell_row
x, y

CH7 Table 5: HDF5 Soil Moisture data

3.6.3 Drought Monitor Algorithm

We follow these general steps to create algorithm for Drought Monitor:

1.Data Preprocessing:

Import the necessary Python libraries, including NumPy, Pandas, xarray, and any specific libraries for working with the chosen data format (e.g., netCDF4).
Read the SMAP L4 data files using the appropriate functions provided by the selected libraries.

2. Calculation of Drought Indices:

Choose a suitable drought index or indices for your drought monitor. Popular choices include the Standardized Precipitation Index (SPI), Standardized Precipitation Evapotranspiration Index (SPEI), or Soil Moisture Anomaly Index (SMAI).
Calculate the selected drought index using the processed SMAP L4 soil moisture data.

3.Threshold Determination:

Define the thresholds that represent different drought severity levels based on your specific requirements and the chosen drought index. These thresholds can be based on established guidelines or customized for your study area and application.
Categorize the calculated drought index values into drought severity levels based on the defined thresholds.

4.Visualization:

Utilize Python libraries like Matplotlib, Seaborn, or Plotly to create visual representations of the drought monitor.
Generate maps, time series plots, or other visualizations that display the spatial and temporal variations of drought severity levels based on the SMAP data and calculated indices.

5.Interpretation and Analysis:

Analyze the drought monitor outputs to gain insights into the drought patterns, trends, and impacts in your study area.
Conduct further statistical analysis or integrate additional data sources (e.g., precipitation, temperature) to enhance the understanding of drought dynamics and their potential implications.[21][22][23]

3.6.3.1 Sequence to create Drought Monitor [17][24][25]

1.Analysis soil Moisture data for April in years from 2015 to 2022

2.Get SMAP Pixel

3.Make 30 day moving average (to smoothen curve)

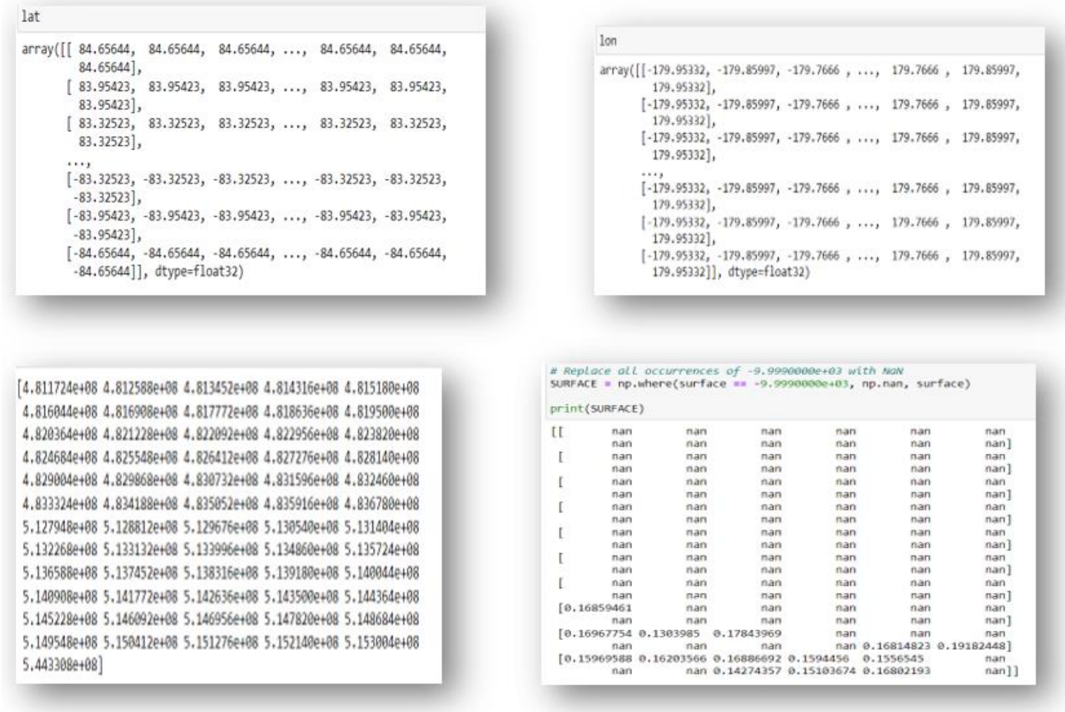
4.Create max/min curve out of all year curves

5.Calculate percentile values for SMI of each pixel

6.Scale depending on percentile

1.Analysis soil Moisture data for April in years from 2015 to 2022

1. 4D-Matrix: 365 days x len(lat) x len(lon) x 5
2. Use Astropy to save into a FITS-File



```
lat
array([[ 84.65644,  84.65644,  84.65644, ...,  84.65644,  84.65644,
         84.65644],
       [ 83.95423,  83.95423,  83.95423, ...,  83.95423,  83.95423,
         83.95423],
       [ 83.32523,  83.32523,  83.32523, ...,  83.32523,  83.32523,
         83.32523],
       ...,
       [-83.32523, -83.32523, -83.32523, ..., -83.32523, -83.32523,
        -83.32523],
       [-83.95423, -83.95423, -83.95423, ..., -83.95423, -83.95423,
        -83.95423],
       [-84.65644, -84.65644, -84.65644, ..., -84.65644, -84.65644,
        -84.65644]], dtype=float32)
```

```
lon
array([[-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332],
       [-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332],
       [-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332],
       ...,
       [-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332],
       [-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332],
       [-179.95332, -179.85997, -179.7666 , ...,  179.7666 ,  179.85997,
          179.95332]], dtype=float32)
```

```
[4.811724e+08 4.812588e+08 4.813452e+08 4.814316e+08 4.815180e+08
 4.816044e+08 4.816908e+08 4.817772e+08 4.818636e+08 4.819500e+08
 4.820364e+08 4.821228e+08 4.822092e+08 4.822956e+08 4.823820e+08
 4.824684e+08 4.825548e+08 4.826412e+08 4.827276e+08 4.828140e+08
 4.829004e+08 4.829868e+08 4.830732e+08 4.831596e+08 4.832460e+08
 4.833324e+08 4.834188e+08 4.835052e+08 4.835916e+08 4.836780e+08
 5.127948e+08 5.128812e+08 5.129676e+08 5.130540e+08 5.131404e+08
 5.132268e+08 5.133132e+08 5.133996e+08 5.134860e+08 5.135724e+08
 5.136588e+08 5.137452e+08 5.138316e+08 5.139180e+08 5.140044e+08
 5.140908e+08 5.141772e+08 5.142636e+08 5.143500e+08 5.144364e+08
 5.145228e+08 5.146092e+08 5.146956e+08 5.147820e+08 5.148684e+08
 5.149548e+08 5.150412e+08 5.151276e+08 5.152140e+08 5.153004e+08
 5.443308e+08]
```

```
# Replace all occurrences of -9.9990000e+03 with NaN
SURFACE = np.where(surface == -9.9990000e+03, np.nan, surface)

print(SURFACE)
[[        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [        nan        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [0.16859461        nan        nan        nan        nan        nan
          nan        nan        nan        nan        nan        nan]
 [0.16967754 0.1303985  0.17843969        nan        nan        nan
          nan        nan        nan        nan 0.16814823 0.19182448]
 [0.15969588 0.16203566 0.16886692 0.1594456  0.1556545         nan
          nan        nan 0.14274357 0.15103674 0.16802193        nan]]
```

CH7 Figure 13

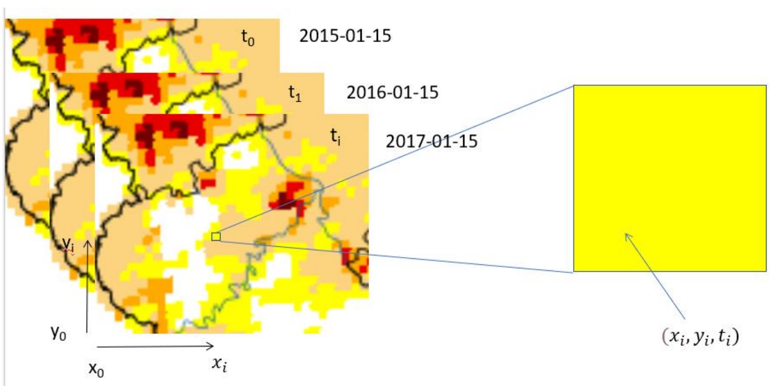
For our Drought Monitor, we will consider `Nan` as the standard value for sea and oceans, and we will only focus on soil moisture values on land.

Additionally, the time values in the HDF5 dataset are represented as Julian Dates.

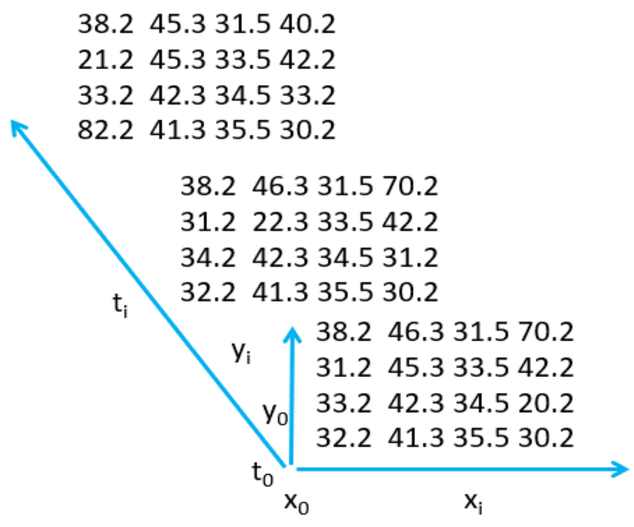
2.Get SMAP Pixel

Create a function to get pixel according to longitude and latitude.

1. **x** longitude
2. **y** latitude
3. **t** time



CH7 Figure 14: SMAP pixel



CH7 Figure 15

3.Make 30 day moving average (to smoothen curve)

We need to calculate the moving average for soil moisture values to get drought indices.

Average value = 365 days for year **x** 30 days for month **x** 24 hours **x** 60 min **x** 60 sec

```
surface_mov_avr[time_idx] = np. average(surface[time_start_idx:(time_end_idx+1)], axis=0)
```

4.Create max/min curve out of all year curves

By using the maximum and minimum values of moving average for soil moisture, we can calculate the corresponding maximum and minimum values of drought indices.

Then save all max and min values in csv file

5.Calculate percentile values for SMI of each pixel

For one month in all years we just calculate min and the first quartile for the reference period from **2015 to 2022**

```python
fifteen_april_time_array = []
i = 0
for x in range(2015,2023):
    print(str(time_specific))
    print(getClosestNeighbour((time_specific+i*(60*60*24*365)),time))
    fifteen_april_time_array.append(getClosestNeighbour(time_specific+i*(60*60*24*365),time))
    i = i+1

# Calculate percentile values for each pixel in lat x lon matrix
Surface_quart = np.percentile(surface_mov_avr[fifteen_april_time_array], 0.25, axis=0)
Surface_min = np.percentile(surface_mov_avr[fifteen_april_time_array], 0.0, axis=0)
Surface_quart.append(np.percentile('Surface_Quart', surface_mov_avr[time_idx][lat_idx][lon_idx], 0.25))
Surface_min.append(np.percentile('Surface_min',surface_mov_avr[time_idx][lat_idx][lon_idx], 0.0))



Surface_quart = np.percentile(surface_mov_avr[fifteen_april_time_array], 0.25, axis=0)
Surface_min = np.percentile(surface_mov_avr[fifteen_april_time_array], 0.0, axis=0)
Surface_quart = np.concatenate([Surface_quart, [np.percentile(surface_mov_avr[time_idx][lat_idx][lon_idx], 0.25)]])
Surface_min = np.concatenate([Surface_min, [np.percentile(surface_mov_avr[time_idx][lat_idx][lon_idx], 0.0)]])
```

CH7 Figure 16

6.Scale depending on percentile

To determine the scaling factors, we will compare the soil moisture values with the first quart values.

scale_factors = [0, 0.2, 0.6, 0.8,0.92]

1 Dry: $SM < SM_{quart} – 0 \times (SM_{quart} - SM_{min})$
2 Moderate Drought: $SM < SM_{quart} – 0.2 \times (SM_{quart} - SM_{min})$
3 Severe Drought: $SM < SM_{quart} – 0.6 \times (SM_{quart} - SM_{min})$
4 Extreme Drought: $SM < SM_{quart} – 0.8 \times (SM_{quart} - SM_{min})$
5 Exceptional Drought: $SM < SM_{quart} – 0.92 \times (SM_{quart} - SM_{min})$

CH7 Table 6 Drought Scale

3.6.3.2 Plotting Drought Monitor

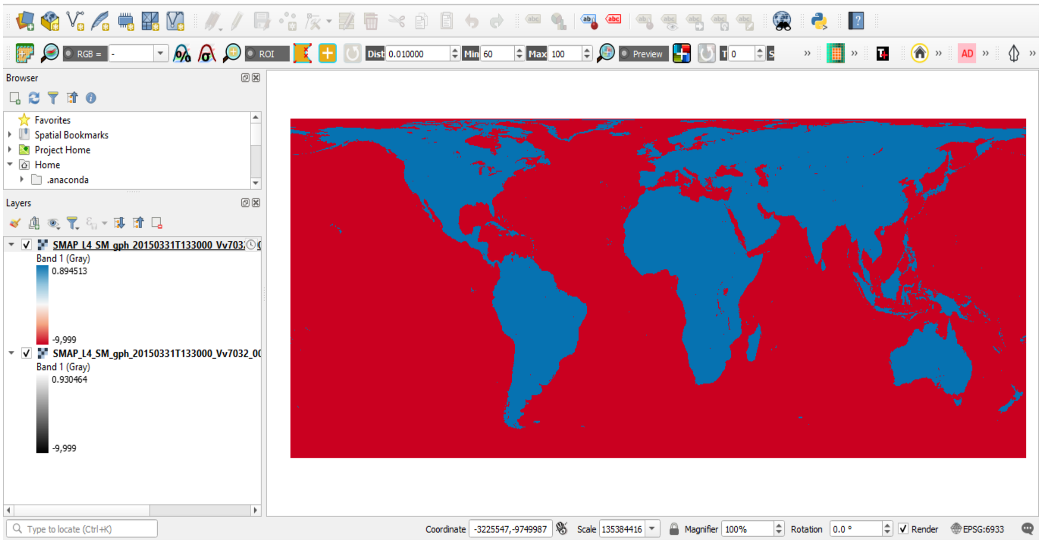
Once we have obtained the soil moisture (SM) values, as well as the minimum, maximum, and quartile values, the next step is to compare these values with established drought scales. This allows us to assess the severity of drought conditions in a given region.

After determining the drought severity, the final step is to create a drought monitor. This is typically done by plotting the drought scale on a graph, with longitude on the x-axis and latitude on the y-axis. This visual representation allows us to easily identify areas that are experiencing drought conditions and to track changes in drought severity over time.

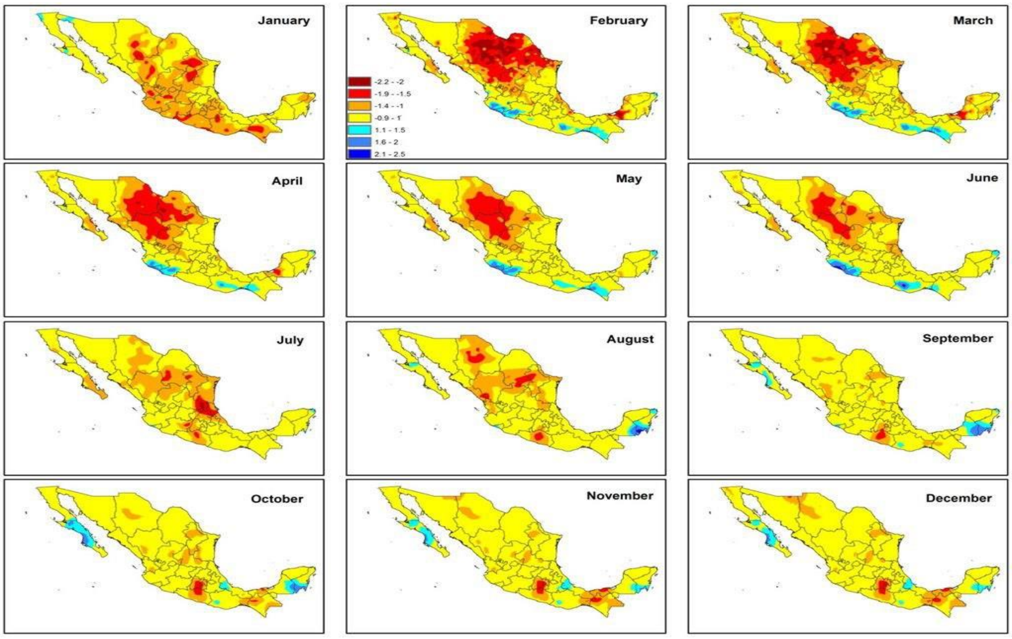
4.Results

4.1 Soil Moisture data analysis with SMAP

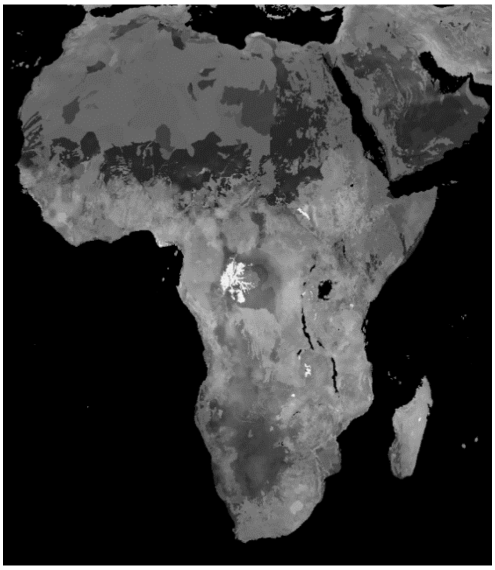
We were able to analysis Geo tiff data of Soil Moisture on QGIS software to see values of soil moisture as color.



CH7 Figure 17: SMAP data analysis on Qgis.

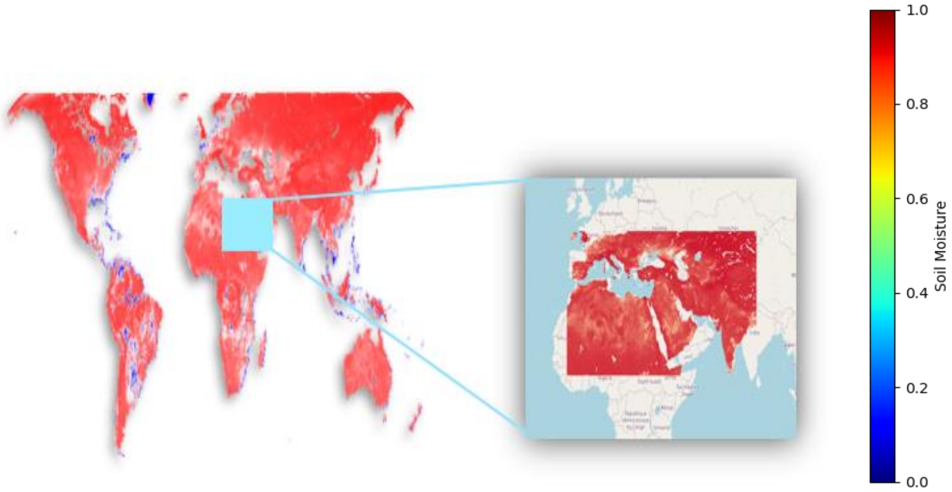


CH7 Figure 18: SMAP data visualize.

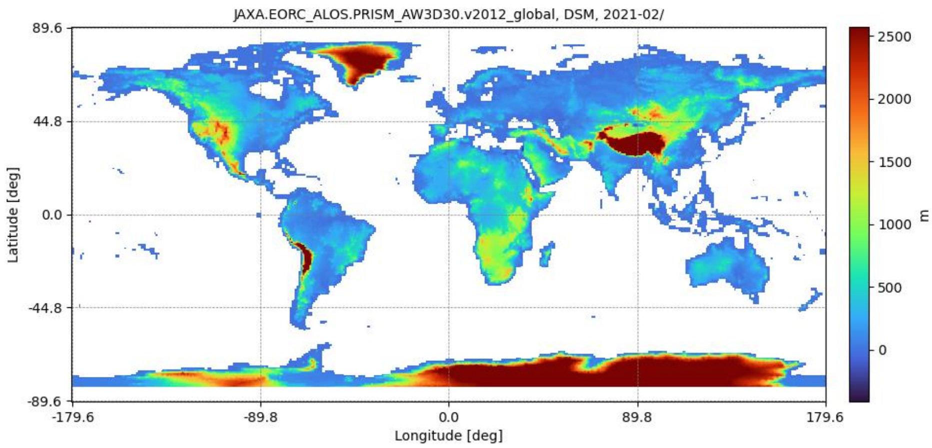


CH7 Figure 19: SMAP analysis in 6-2023

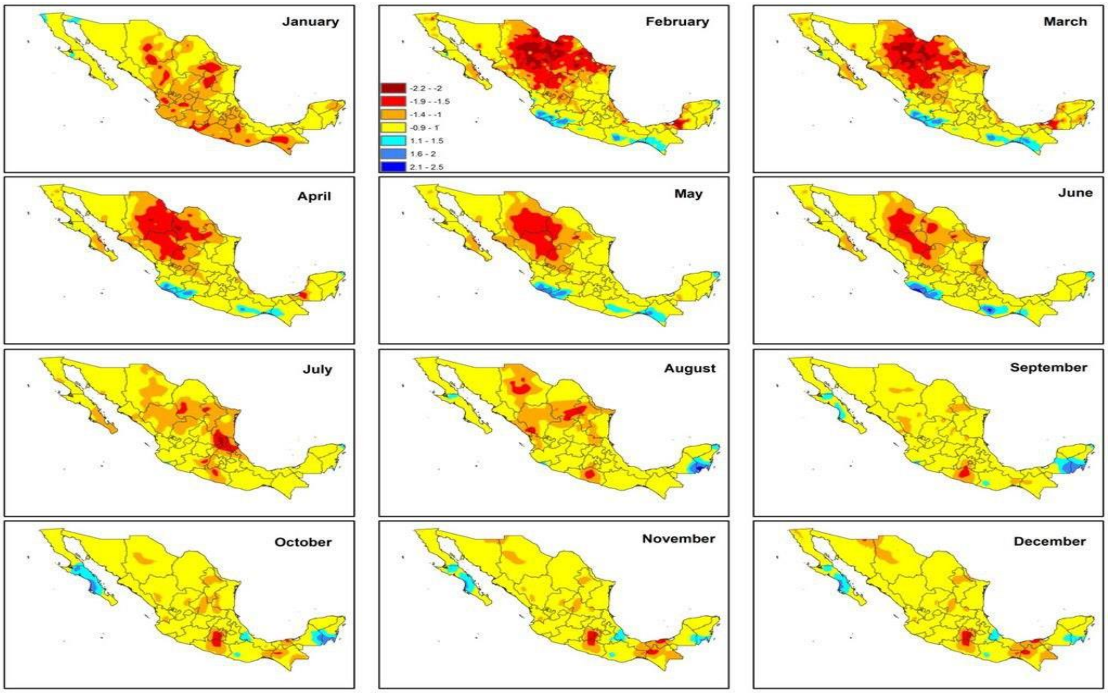
4.2 Soil Moisture data analysis with JAXA



CH 7 Figure 20: JAXA soil moisture data analysis using Qgis.

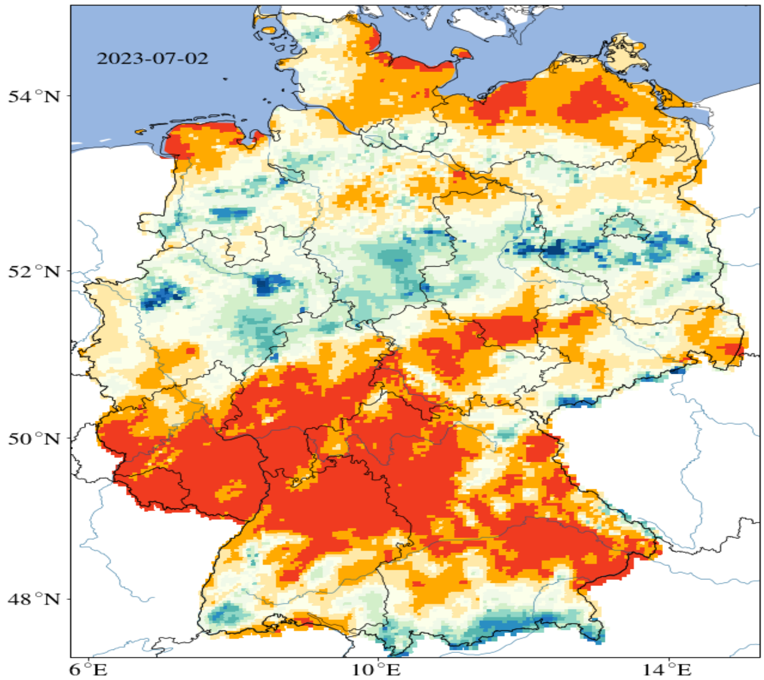


CH7 Figure 21: JAXA soil moisture analysis using python.

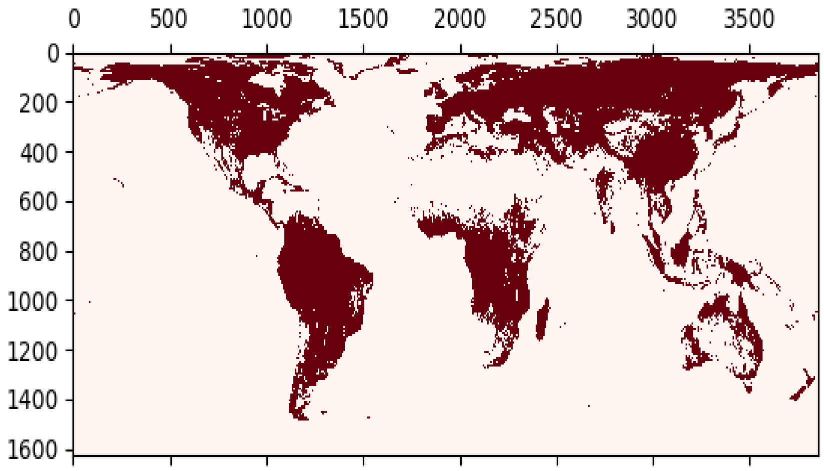


CH7 Figure 22

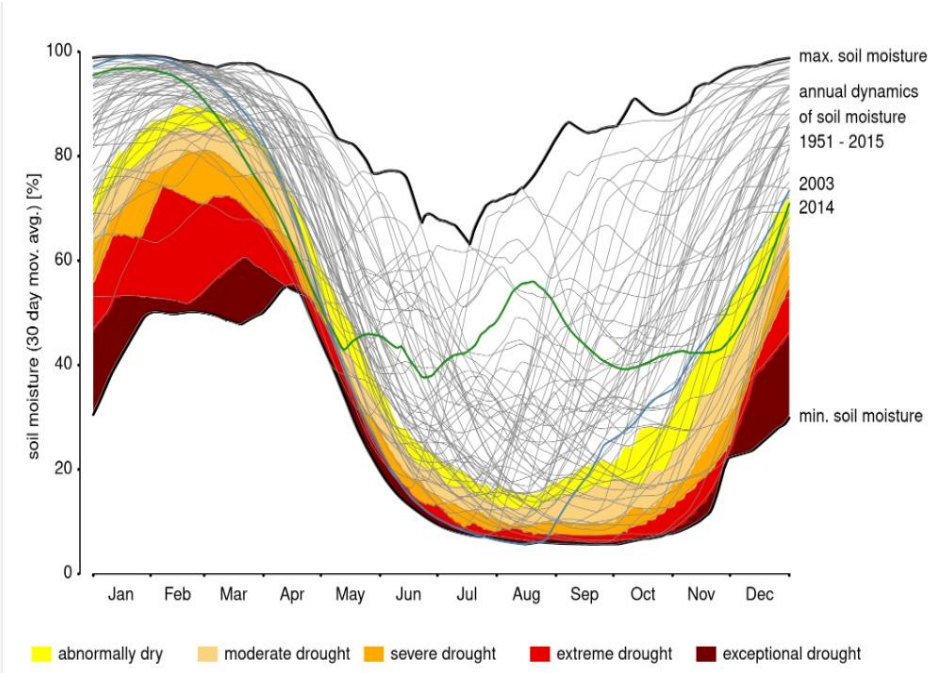
4.3 Drought Monitor



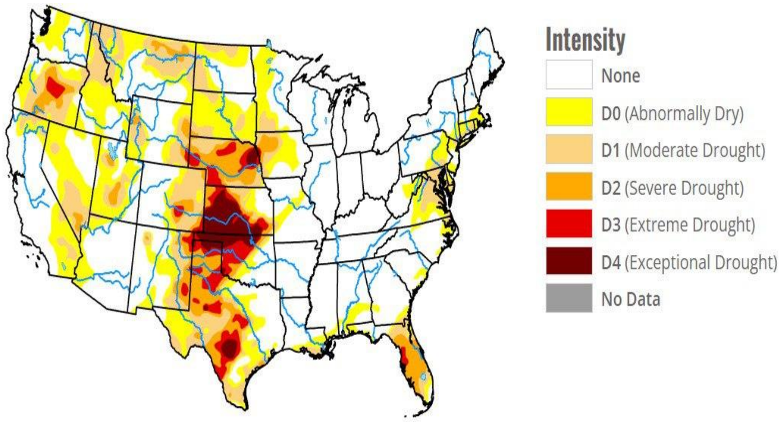
CH 7 Figure 23 Drought Monitoring in 2022 Germany



CH7 Figure 24 Worldwide drought Monitoring



CH7 Figure 25 Vol. Soil Moisture values at a location in Germany, with drought percentile.



CH7 Figure 26: DROUGHT MONITOR WITH DROUGHT SCALE

5. Conclusion

The SMAP L4 soil moisture analysis provides valuable insights into soil moisture patterns, benefiting various scientific disciplines and practical applications. Analysis improves our understanding of soil moisture variability, which is vital for studying hydrological processes, assessing drought conditions, and managing water resources effectively. The analysis also enhances agricultural applications by providing accurate soil moisture information, improving crop yields, and water efficiency. Incorporating SMAP L4 soil moisture data into weather and climate models helps enhance the accuracy of forecasts, especially in areas prone to droughts, floods, and extreme weather events. The analysis contributes to monitoring vegetation growth, assessing soil moisture impacts on ecosystems, and studying interactions between land surface processes and climate change. SMAP L4 soil moisture analysis serves as a reference dataset for validating and calibrating other satellite missions and remote sensing products, ensuring data consistency and accuracy across different platforms, and improving the overall quality of soil moisture observations. Overall, SMAP L4 soil moisture analysis plays a crucial role in improving our understanding of the Earth's water cycle, supporting sustainable water resource management, agriculture, weather prediction, and climate studies. The creation of a drought monitor from SMAP data can be a valuable tool for assessing and managing drought conditions. SMAP provides high-resolution, global soil moisture data that is crucial for monitoring drought. Through analyzing this data, scientists and policymakers can obtain accurate and up-to-date information about soil moisture conditions, enabling them to identify regions experiencing drought, assess the severity and extent of drought conditions, and make informed decisions regarding water resource management and agricultural practices. Creating a drought monitor involves processing and analyzing satellite imagery and integrating it with other relevant data sources to allow for a more comprehensive understanding of drought conditions. The monitor can be developed using various techniques, including statistical models, machine learning algorithms, and remote sensing techniques, to classify different drought categories based on soil moisture data and associated variables. Implementing a drought monitor can improve early warning systems, aid in drought forecasting and prediction, and assist farmers, water managers, and policymakers in making informed decisions related to irrigation planning, water allocation, and drought response strategies. Overall, leveraging SMAP data for drought monitoring can enhance our ability to assess, understand, and respond to drought conditions, contributing to better drought management and water resource planning.

8.OIL STORAGE DETECTION

Abstract

This thesis project aims to develop a remote sensing solution that utilizes innovative machine learning and AI algorithms to process, analyze, and interpret satellite data. The primary objective is to create a highly efficient AI model capable of detecting and classifying specific features, including trees, buildings, streets, farms, and oil fields. The research will utilize high-resolution training datasets from Google Earth, which have been labeled to indicate the location of bounding boxes around each oil storage. The project will focus on oil storage detection, given its critical role in infrastructure management, impacting multiple sectors, such as industry and energy. The AI model will be trained to estimate the location and size of oil storage, which will help improve their management, prevent any potential risks, and enhance monitoring capabilities. The project will utilize the state-of-the-art YoloV5 algorithm, in conjunction with tools and workflows on AWS, to train, validate, and assess the DL model for object detection. In this thesis, we utilized the Segment Anything Model (SAM) to convert object detection results into instance segmentation results, with the aim of improving the accuracy and quality of the segmentation outputs. The proposed solution has significant potential to enhance infrastructure management, particularly in monitoring oil tanks, thereby contributing to the overall efficiency and safety of the infrastructure. The research will involve developing and testing various models using different parameters and optimization techniques to ensure high accuracy and reliability. The project will also evaluate the feasibility of the proposed solution by comparing it to other existing approaches and solutions. Furthermore, the research will explore the potential applications of the developed model in other fields, such as agriculture, urban planning, and environmental monitoring. The project will also discuss the ethical and social implications of using AI in remote sensing and infrastructure management. This thesis project seeks to develop a novel and innovative solution for remote sensing and infrastructure management, leveraging state-of-the-art machine learning and AI techniques to improve efficiency and safety while contributing to sustainable development.

1- Introduction

Oil storage is a critical component of the energy industry and plays a vital role in the global economy. Effective management of oil tanks is essential for ensuring a reliable supply of oil and mitigating potential risks, such as spills, which can harm the environment and disrupt the economy. Therefore, accurate and efficient detection, classification, and monitoring of oil storage are crucial for infrastructure management. Traditional remote sensing methods for oil tank detection require manual interpretation and analysis of satellite imagery, which can be time-consuming and prone to errors. To address these challenges, machine learning and artificial intelligence (AI) algorithms have been employed to automate the process of oil tank detection and classification. This thesis aims to prototype a remote sensing solution that leverages machine learning and AI algorithms to process, analyze, and interpret satellite data for infrastructure management, with a focus on oil storage detection. The project will train an AI model to detect and classify specific features, such as oil fields, using labeled training datasets obtained from Google Earth. The proposed solution will utilize the state-of-the-art YoloV5 library, along with tools and workflows on AWS, to train, validate, and assess the DL model for object detection, with a focus on oil tank monitoring and reducing potential risks, and using Segment Anything Model (SAM) for converting object detection results into instance segmentation results. The research will involve developing and testing various models using different parameters and optimization techniques to ensure high accuracy and reliability. The thesis project aims to develop a highly accurate and efficient AI model that can enhance infrastructure management, particularly in the monitoring of oil tanks, thus mitigating potential risks, such as spills. The research will evaluate the feasibility of the proposed solution by comparing it to other existing approaches and solutions. Furthermore, the research will explore the potential applications of the developed model in other fields, such as agriculture, urban planning, and environmental monitoring. The project aims to develop an innovative solution for remote sensing and infrastructure management, leveraging innovative machine learning and AI techniques to improve efficiency and safety while contributing to sustainable development. The research will focus on oil storage detection as a critical use case and will utilize labeled training datasets sourced from Google Earth.

2- Literature Review

Kasper-Eulaers et al. (2021) [**37**,**38**] studied how YOLOv5 can be implemented to detect heavy goods vehicles at rest areas during winter to allow for the real-time prediction of parking spot occupancy. The model was trained using Google Colaboratory (Colab), which provides free access to powerful GPUs and requires no configuration. A notebook was developed by Roboflow.ai which is based on YOLOv5 and uses pre-trained COCO weight. The model improved swiftly in terms of precision, recall, and mean average precision before overfitting after about 150 epochs. The box, objectness and classification losses of the validation data also showed a rapid decline until around epoch 15. Results show that the trained algorithm can detect the front cabin of heavy goods vehicles with high confidence, while detecting the rear seems more difficult, especially when located far away from the camera. Malta et al. (2021) [**37**,**39**] proposed a model of a task assistant based on a deep learning neural network. A YOLOv5 network was used for recognizing some of the constituent parts of an automobile. The dataset created consisted of 582 images taken from three videos with similar lighting conditions, where it was possible to identify a total of eight different types of parts: oil dipstick; battery; engine oil reservoir; wiper water tank; air filter; brakes fluid reservoir; coolant reservoir; and power steering reservoir. The images taken from each frame were converted to a (416 × 416) format, which is the format that the chosen architecture needs to use as input. The hardware used during development included computers, for running software, and cell phones, for capturing videos and pictures. The object detection model was trained using laptop computer with access to a Google Colab virtual machine. The precision obtained for the two models (YOLOv5s and YOLOv5m) was in line with that obtained by other authors for similar problems. YOLOv5s demonstrated to be capable of identifying eight different mechanical parts in a car engine with high precision and recall always above 96.8% in the test sets, which, compared to the larger model, has almost the same results. Results proved that the network is good and fast enough to be applied to the task of assisting in recognizing constituent parts of an automobile.

Wan et al. (2021) [**37**, **40**] proposed a YOLOv5 model based on a self-attention mechanism for polyp target detection. Mosaic method was used in the data preprocessing stage to enhance the amount of training data in the dataset, Cross Stage Partial Networks (CSPNet) were used as the backbone network to extract the information features in the image, which solved the problem of gradient disappearance, and the feature pyramid architecture with attention mechanisms was used to enhance the detection performance of varying-size polyps. The proposed method was trained by stochastic gradient descent (SGD) and backpropagation in an end-to-end way on a cloud-computing platform configured with eight 16 GB GPUs, a 16-core CPU, and a 64 GB memory. YOLOv5 used spatial pyramid pooling (SPP) to enhance the model's detection of objects with different scales, Path aggregation network (PANET) as the neck for feature aggregation and new Feature Pyramid Networks (FPN) structure that enhanced the bottom-up path, which improved the propagation of low-level features. The author's method achieved excellent performance. In the Kvasir-SEG data set, the precision was 0.915, the recall rate was 0.899 and the Fscore was 0.907. In the WCY data set, the precision was 0.913, the recall was 0.921 and the F-score was 0.917. Specifically, this method used full-image information when predicting the target window using each network, which greatly reduced the false positive rate.

Yao et al. (2021) [**37**, **41**] developed a defect detection model based on YOLOv5, which can detect defects accurately, and at a fast speed. A small object detection layer was added to improve the model's ability to detect small defects. Squeeze-and-Excitation (SE) Layer and the loss function complete intersection over union (CIoU) were introduced to make the regression more accurate. The model was trained based on transfer learning and used the Cosine Annealing algorithm to improve the effect. The mAP@0.5 of YOLOv5 reached 94.7%, which was an improvement of nearly 9%, compared to the original algorithm.

The paper [**5**], based on the YOLO v5 object detection algorithm, a power plant oil reservoir oil leakage detection algorithm based on the CBAM-YOLO v5 model is proposed, and CBAM is introduced [**35**]. The attention mechanism makes the model focus more on the extraction of pipeline leakage features, weakens the influence of the complex background of the power plant on the detection results, uses Adaptively Spatial Feature Fusion to stitch the features of different levels after convolution, and suppresses inconsistencies by learning the contradictory information of spatial filters, improves the scale invariance of features and reduces the computational overhead in the reasoning process. Based on meeting the requirements of real-time detection, the algorithm proposed in the paper [**5**] can accurately locate the location of the oil leakage, and has a good accuracy, realizing the automatic detection and early warning of the on-site oil leakage. In summary, this paper improves based on the single-stage target algorithm YOLOv5s that considers the recognition speed and detection accuracy, and the YOLO v5s model is more lightweight, and the number of parameters of the YOLO v5s model selected in this paper is only 7.2M. The input image size selection is more flexible, YOLO v5 can adaptively scale the input image to a unified standard size, to adapt to the complex environmental factors of the power plant oil depot, by integrating the CBAM attention mechanism, reassign the appropriate weight coefficient to the features of different resolutions, enhance the useful features in the image, suppress useless features, and make the model pay corresponding attention to the areas where oil leakage may be in the image The general splicing operation is replaced by Adaptively Spatial Feature Fusion, and the features of different levels are adaptively fused to better retain the useful information in the features. The main contributions of this research are:

- An oil leakage detection algorithm based on the improved YOLO v5 is proposed for power plant oil depot pipelines. By integrating the CBAM attention mechanism and Adaptively Spatial Feature Fusion, the model focuses more on extracting oil leakage features and improves the scale invariance of features.
- A self-built oil leakage dataset from power plant pipelines is constructed. Experiments show that the proposed algorithm has higher accuracy and a lower false detection rate than the original YOLO v5 algorithm.

The proposed algorithm shows strong generalization performance in the complex environment of power plants. It can detect oil leakage at different distances and angles accurately. This paper is directly relevant to this thesis as it also explores using the YOLOv5 framework for object detection.

The paper [**36**], the single-stage target detection algorithm based on deep learning is combined with the UAV patrol line of natural gas pipelines, and an algorithm model based on improved YOLOv5 is proposed. The Recurrent Criss-Cross Attention module is added to the backbone network to improve the feature extraction capability. The feature fusion module uses the TridentNet network to improve the performance of the multi-scale feature fusion of the model. WBF post-processing is used to replace the original NMS, improving the prediction algorithm's regression accuracy. Experiments show that the improvement has effectively improved the detection effect of third-party intrusion of natural gas pipelines. In the third-party intrusion detection task of natural gas pipelines, fusing features of different scales is an important means to improve the performance. The low-level features have higher resolution and often contain more construction site location and detailed information, but due to fewer convolutions, they have lower semantics and more noise. High-level features have stronger semantic information, but low resolution and poor perception of details. YOLOv5 adopts the structure of FPN+PAN to realize feature extraction. The network draws on the pyramid structure to combine resolution information and semantic information to overcome the problems caused by different layers and different scales. But there is still room for improvement in accuracy. This paper uses Trident Block for feature fusion. This module combines the image pyramid and feature pyramid at the same time and uses the parameter sharing method to propose three branches for training, but only one branch is used for testing, to ensure that there will be no additional parameters and calculations during forward reasoning volume increase. In summary, these two papers present complementary applications of deep learning for oil spill monitoring through image classification and localization. The proposed work in this thesis can build upon these papers by adapting YOLOv5 for a comprehensive solution to detect oil storage tanks.

3- Methodology

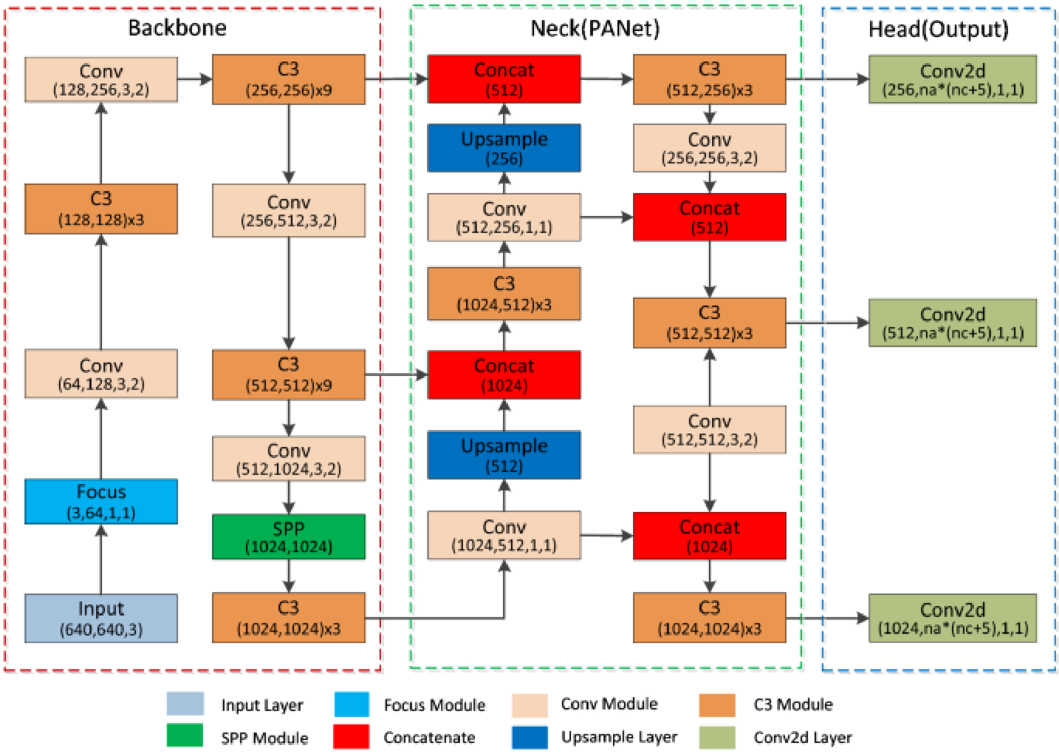
This thesis is divided into two parts, Part one: Object Detection and Part two: Transformation from object detection to instance segmentation using SAM.

<u>Methodology for part one:</u>

3.1 Yolov5 Architecture [5]

The YOLOv5 Algorithm is a state-of-the-art object detection algorithm that has gained significant attention in the computer vision community due to its high detection accuracy and fast inference speed. This algorithm is a one-stage object detection algorithm that directly predicts the location and category of an object after extracting its features. It has been shown to outperform other popular algorithms such as Faster R-CNN and Mask R-CNN in terms of detection speed and accuracy. The YOLOv5 Algorithm is an improvement over its predecessors YOLOv4, YOLOv3, and YOLOv2. In YOLOv4, the backbone network was modified to CSPDarknet53, the activation function was replaced with Mish, and a PANet structure based on the FPN from YOLOv3 was added. YOLOv5 has further improved based on YOLOv4, with faster detection speed and reduced model size, making it more suitable for real-world applications. The YOLOv5 algorithm has four models: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. These models differ in depth and width, with increasing feature extraction ability. In this project, the largest YOLOv5x network was selected as the research object as it is intended to be deployed on Websites. The algorithm was further improved based on the specific needs of the study, and a model for the detection of oil storage tanks, floating head tanks, and tank clusters was established.

The YOLOv5 Algorithm is composed of three parts: backbone, neck, and head. The backbone consists of a series of convolutional neural networks (CNNs) used to extract image features. The backbone includes focus, C3, and SPP modules. The focus module slices the images and splices them into the channel dimension to integrate the width and height information into the channel dimension, which can effectively improve the speed of feature extraction. The C3 module is improved from the structure of the cross-stage partial (CSP) connections, having one less convolution layer and changing the activation function, and its main function is to extract features from images and reduce the repetition of gradient information. The spatial pyramid pooling (SPP) module respectively uses three pooling kernels of sizes 5, 9, and 13 to perform max-pooling operations on the images. This module can increase the receptive field of the network and obtain features of different scales. The neck is the feature fusion network of the model, where feature pyramid networks (FPNs) and path aggregation network (PANet) are adopted. The structure of the FPNs transmits semantic information from the top down, while the PANet additionally transmits location information from the bottom up based on the FPNs. The head is the prediction network of the model and, through convolution operations, three groups of feature vectors containing the categories prediction boxes, confidence, and coordinate position are output, which predicts at scales of (80 × 80), (40 × 40), and (20 × 20), respectively.



CH 8 Figure1 YOLOv5 Architecture

3.2 The key features of the YOLO v5 object detection algorithm [5]:

3.2.1) Mosaic data enhancement

At the input end of YOLO v5, mosaic data enhancement operation is used, which divides the input pictures into four groups, and stitches the four pictures in each group according to the way of scaling, then cropping and random arrangement, and finally obtains a new picture and passes it into the neural network to learn, which greatly enriches the background of detecting objects, reduces the algorithm's dependence on the match size and improves the training speed and network accuracy of the model.

3.2.2) Adaptive anchor frame calculation

In the YOLO series of algorithms, for different data sets, it is necessary to set a specific length and width anchor point frame, and the setting of the initial anchor point frame is also a key part of the initial stage of network training. In YOLO v5, the algorithm can adaptively calculate the optimal anchor box based on the name of the dataset at each training.

3.2.3) Adaptive image zoom

In the object detection algorithm, the length and width of different images are different, so the more common way is to uniformly scale the original image to a standard size and then feed it into the detection network. There are many problems with the traditional scaling method, if too much padding after scaling, there will be a lot of information redundancy, which will affect the inference speed of the entire algorithm. Therefore, the YOLO v5 algorithm proposes an adaptive scaling method with the least black edges, first calculating the scale according to the size of the original picture and the size of the picture entered the network, then calculating the size of the scaled picture according to the original picture size and scale, and finally calculating the black edge fill value for filling. This adaptive image scaling can improve the inference speed of the algorithm.

3.2.4) Focus operation

In the backbone segment, the YOLO v5 algorithm uses the Focus operation to slice the picture, obtains a value for every other pixel in a picture (similar to proximity downsampling), and finally obtains four similar images, and then the obtained new image is convoluted to obtain a feature map that is doubled and subsampled, and there is no information loss, the height and width information is concentrated in the channel space during the whole slicing process, and the input channel is expanded by 4 times. That is, the input channel has changed from the original R, G, B three channels to 12 channels.

3.2.5) Cross Stage Partial Network (CSP network)

The CSP network splits the input features into two parts, one part is convoluted, and the other part is merged with the results of the previous convolution operation (Concat), which solves the gradient information duplication problem of network optimization in the large convolutional neural network architecture, integrates the gradient change into the feature graph, and takes into account the model inference speed and accuracy while reducing the size of the model by reusing the feature map.

3.2.6) Path-Aggregation Network (PANet)

PANet is used to aggregate features, and the feature extractor of the network uses a new FPN structure that enhances the bottom-up path to improve the propagation of low-level features so that the low-level positioning information can also be enhanced to the entire feature extraction network. By adjusting the depth multiple and width multiple parameters, YOLO v5 implements YOLO v5s, YOLO v5m, YOLO v5 l, and YOLO v5x four structures, these four structures networks gradually deepened and widened, generalization ability, feature extraction ability, feature fusion ability, and AP accuracy continue to improve, but with the follow-up is that the model training speed will slow down with the increase or decrease of model complexity. YOLOv5 model achieves high accuracy and speed through a combination of different components, including the activation function, optimization function, cost function or loss function, Weights, biases, parameters, gradients, and final model summary. This thesis analyzes each component and its significance in the YOLOv5 model. The thesis also evaluates the impact of different components on the performance of the YOLOv5 model through experimentation and analysis.

3.3 Activation Function [42]

The activation function is a critical component of neural networks used in the YOLOv5 model. The activation function determines whether a neuron should be activated or not. The YOLOv5 model uses different activation functions, including ReLU, LeakyReLU, and Sigmoid. The ReLU activation function is the most widely used activation function in the YOLOv5 model due to its simplicity and speed. However, the LeakyReLU activation function performs better in detecting objects with low contrast, while the Sigmoid activation function is useful in detecting objects with non-linear boundaries.

3.4 Optimization Function [42]

The optimization function is another essential component of neural networks used in the YOLOv5 model. The optimization function is responsible for minimizing the error between the predicted output and the actual output. The YOLOv5 model uses different optimization functions, including Adam and SGD. The Adam optimization function is the most widely used optimization function in the YOLOv5 model due to its ability to adapt to different learning rates. However, the SGD optimization function performs better in detecting objects with complex boundaries.

3.5 Cost Function or Loss Function [42]

The cost function or loss function is a critical component of neural networks used in the YOLOv5 model. The cost function or loss function measures the error between the predicted output and the actual output. The YOLOv5 model uses different cost functions, including Mean Squared Error (MSE) and Binary Cross-Entropy (BCE). The MSE cost function is the most widely used cost function in the YOLOv5 model due to its simplicity. However, the BCE cost function performs better in detecting objects with complex boundaries.

3.6 Weights, Biases, Parameters, Gradients, and Final Model Summary [42]

weights, biases, parameters, and gradients are essential components of neural networks used in the YOLOv5 model. These components are responsible for adjusting the model's output to match the actual output. The final model summary provides an overview of the YOLOv5 model's architecture. The YOLOv5 model uses different Weights, biases, parameters, and gradients to adjust the model's output. The final model summary provides an overview of the YOLOv5 model's architecture, including the number of layers, filters, and parameters.

3.7 Results of Output [42]

The YOLOv5 model achieves high accuracy and speed in detecting objects in real-time. It outperforms other object detection models, including YOLOv4 and Faster R-CNN, in terms of accuracy and speed. The results demonstrate the importance of each component in achieving high accuracy and speed in object detection. The performance of the model is evaluated using metrics such as precision, recall, F1 score, and mean average precision (mAP).

3.7.1 Precision measures the proportion of true positive detections among all the objects detected.

3.7.2 Recall measures the proportion of true positive detections among all the objects that should have been detected.

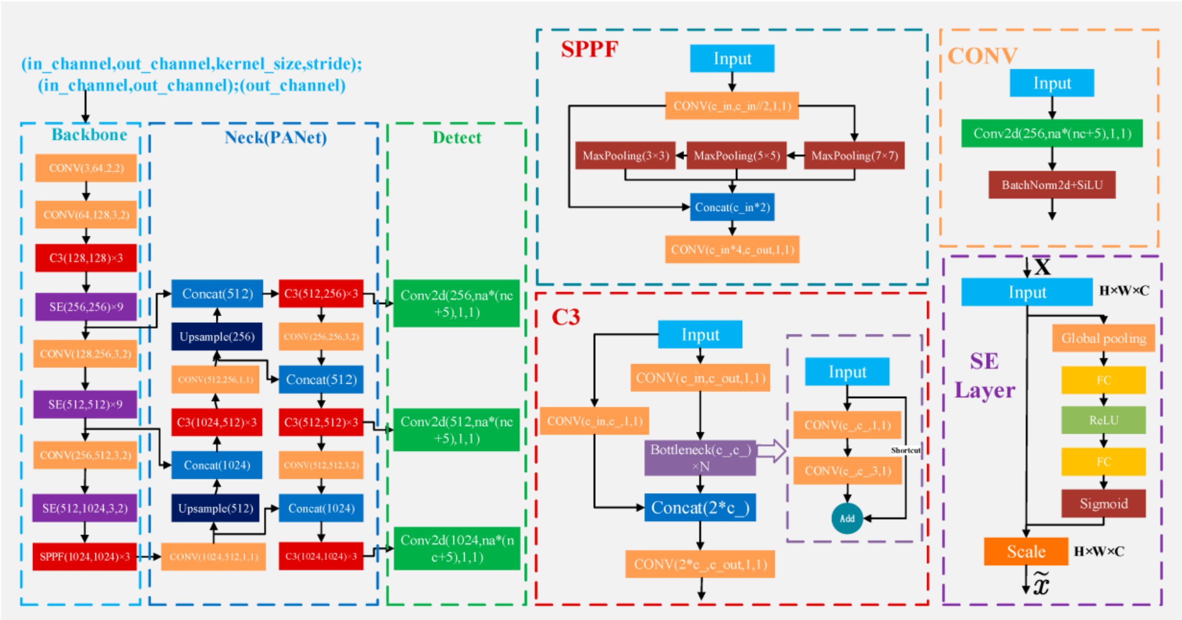
3.7.3 F1 score is the harmonic means of precision and recall and provides a balanced measure of the model's performance.

3.7.4 Mean average precision (mAP) is a commonly used metric in object detection tasks that measures the average precision across different recall levels. It is computed by taking the area under the precision-recall curve. In the experiments conducted for this thesis, the YOLOv5 model achieved high precision, recall, F1 score, and mAP across different datasets and object classes. The model's performance was consistently high across different activation functions, optimization functions, and cost functions. However, it was observed that the performance of the model was affected by the size and complexity of the objects being detected. Objects with small sizes or complex boundaries were more challenging to detect, leading to lower precision and recall. Overall, the results demonstrate that the YOLOv5 model is a highly effective object detection algorithm that achieves high accuracy and speed in real-time. The model's performance can be further improved by optimizing the detection of small objects and complex boundaries.

3.8 Architecture of YOLOv5x [43]

Among the four models in the YOLOv5 family, the YOLOv5x model is the most advanced and powerful model, which has been designed to provide high accuracy object detection in complex scenes. The YOLOv5x model is equipped with a CSPDarknet53 backbone network, which is an improved version of the CSPDarknet network used in YOLOv4. The CSPDarknet53 network has a hybrid residual and dense connection structure that reduces the number of parameters while improving the feature extraction ability. This network architecture enables the YOLOv5x model to extract more detailed features from the input images.

The YOLOv5x model employs advanced feature extraction techniques, including focus, C3, and SPP modules. The focus module slices the images and splices them into the channel dimension, effectively enhancing the speed of feature extraction. The C3 module uses the structure of cross-stage partial (CSP) connections to extract features from the images and reduce the repetition of gradient information. The spatial pyramid pooling (SPP) module uses three pooling kernels of size 5, 9, and 13 to perform max-pooling operations on the images, which can increase the receptive field of the network and obtain features of different scales. The neck of the YOLOv5x model uses feature fusion techniques to combine features from different scales and levels. The model uses PANet, which is an extension of the feature pyramid network (FPN) used in other YOLOv5 models. PANet adds a bottom-up path to the FPN, enabling it to combine features from different spatial resolutions and improve the localization accuracy of small objects. The head of the YOLOv5x model uses a modified YOLOv4 head with more convolutional layers and filters. The model predicts objects at three different scales: 320x320, 640x640, and 1280x1280. The model also uses anchor boxes with different aspect ratios and scales to improve the detection accuracy of objects with different shapes and sizes. Despite having many parameters, with a total of 176 million, the YOLOv5x model can still achieve real-time detection on a single GPU, thanks to its efficient implementation and parallel processing capabilities. The model's high accuracy and real-time performance make it suitable for applications such as autonomous driving and video surveillance. In conclusion, the YOLOv5x model is a powerful object detection algorithm that offers high detection accuracy and fast inference speed in complex scenes. Its advanced feature extraction and fusion techniques, including the CSPDarknet53 backbone network, focus, C3, SPP modules, and PANet neck, make it a state-of-the-art model for object detection. Its ability to achieve real-time performance on a single GPU, despite having many parameters, makes it suitable for a wide range of applications, including autonomous driving, video surveillance, and robotics.

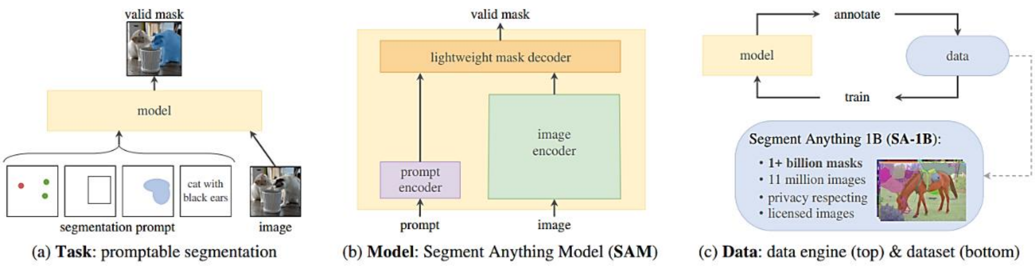


CH 8 Figure 2 YOLOv5x Architecture

Methodology for Part Two: Oil Storage Segmentation Using SAM

3.9 SAM Architecture [8]

Meta AI aimed to build a foundation model for segmentation by introducing three interconnected components: a promptable segmentation task, a segmentation model (SAM) that powers data annotation and enables zero-shot transfer to a range of tasks via prompt engineering, and a data engine for collecting SA-1B, the dataset of over 1 billion masks as illustrated in figure (3).



CH 8 Figure3 Interconnected Components for Segmentation

3.9.1 Segment Anything Task

prediction task is used for foundation model pre-training and to solve diverse downstream tasks via prompt engineering [**9**]. To build a foundation model for segmentation, Meta AI aimed to define a task with analogous capabilities.

Task. They started by translating the idea of a prompt from NLP to segmentation, where a prompt can be a set of foreground/background points, a rough box or mask, free-form text, or, in general, any information indicating what to segment in an image. The promotable segmentation task, then, is to return a valid segmentation mask given any prompt. The requirement of a "valid" mask simply means that even when a prompt is ambiguous and could refer to multiple objects (e.g., recall the shirt vs. person example, and see Fig. [**4**]), the output should be a reasonable mask for at least one of those objects. This requirement is similar to expecting a language model to output a coherent response to an ambiguous prompt. This task was chosen because it leads to a natural pre-training algorithm and a general method for zero-shot transfer to downstream segmentation tasks via prompting.

Pre-training. The promptable segmentation task suggests a natural pre-training algorithm that simulates a sequence of prompts (e.g., points, boxes, masks) for each training sample and compares the model's mask predictions against the ground truth. They adapted this method from interactive segmentation [**11**, **10**], although unlike interactive segmentation whose aim is to eventually predict a valid mask after enough user input, the aim is to always predict a valid mask for any prompt even when the prompt is ambiguous. This ensures that a pre-trained model is effective in use cases that involve ambiguity, including automatic annotation.

Zero-shot transfer. Intuitively, the pre-training task endows the model with the ability to respond appropriately to any prompt at inference time, and thus downstream tasks can be solved by engineering appropriate prompts. For example, if one has a bounding box detector for cats, cat instance segmentation can be solved by providing the detector's box output as a prompt to the model. In general, a wide array of practical segmentation tasks can be cast as prompting.

CH 8 Figure 4: Each column shows 3 valid masks generated by SAM from a single ambiguous point prompt (green circle).

3.9.2. Segment Anything Model

The Segment Anything Model (SAM) is described for promptable segmentation. SAM has three components, illustrated in Fig. 5: an image encoder, a flexible prompt encoder, and a fast mask decoder. We build on Transformer vision models [**12**, **13**, **14**, **15**] with specific tradeoffs for (amortized) real-time performance.

Image encoder. Motivated by scalability and powerful pretraining methods, Meta Researchers use an MAE [**16**] pre-trained Vision Transformer (ViT) [**14**] minimally adapted to process high resolution inputs [**15**]. The image encoder runs once per image and can be applied prior to prompting the model.

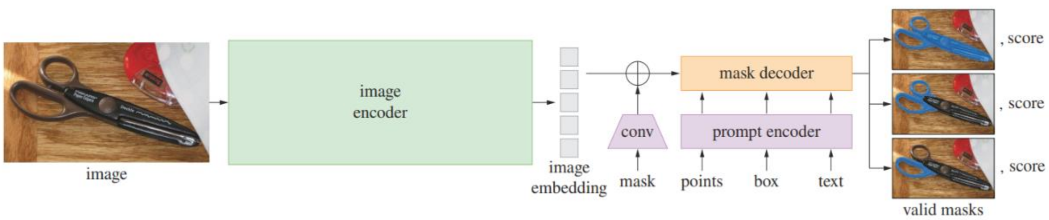
Prompt encoder. They consider two sets of prompts: sparse (points, boxes, text) and dense (masks). They represent points and boxes by positional encodings [**17**] summed with learned embeddings for each prompt type and free-form text with an off-the-shelf text encoder from CLIP [**18**]. Dense prompts (i.e., masks) are embedded using convolutions and summed elementwise with the image embedding.

Mask decoder. The mask decoder efficiently maps the image embedding, prompt embeddings, and an output token to a mask. This design, inspired by [**12**, **13**], employs a modification of a Transformer decoder block [**19**] followed by a dynamic mask prediction head. The modified decoder block uses prompt self-attention and cross-attention in two directions (prompt-to-image embedding and vice-versa) to update all embeddings. After running two blocks, Meta AI Researchers upsample the image embedding and an MLP maps the output token to a dynamic linear classifier, which then computes the mask foreground probability at each image location.

Resolving ambiguity. With one output, the model will average multiple valid masks if given an ambiguous prompt. To address this, Meta Researchers modify the model to predict multiple output masks for a single prompt (see Fig. **3**). Meta AI Researchers found 3 mask outputs are sufficient to address most common cases (nested masks are often at most three deep: whole, part, and subpart). During training, Meta AI Researchers backprop only the minimum loss [**20**, **21**, **22**] over masks. To rank masks, the model predicts a confidence score (i.e., estimated IoU) for each mask.

Efficiency. The overall model design is largely motivated by efficiency. Given a precomputed image embedding, the prompt encoder and mask decoder runs in a web browser, on CPU, in ∼50ms. This runtime performance enables seamless, real-time interactive prompting of the model.

Losses and training. Meta AI Researchers supervise mask prediction with the linear combination of focal loss [**23**] and dice loss [**24**] used in [**12**]. Meta Researchers train for the promptable segmentation task using a mixture of geometric prompts. Following [**25**, **26**], Meta Researchers simulate an interactive setup by randomly sampling prompts in 11 rounds per mask, allowing SAM to integrate seamlessly into the data engine.



CH 8 Figure 5: Segment Anything Model (SAM) overview. A heavyweight image encoder outputs an image embedding that can then be efficiently queried by a variety of input prompts to produce object masks at amortized real-time speed. For ambiguous prompts corresponding to more than one object, SAM can output multiple valid masks and associated confidence scores.

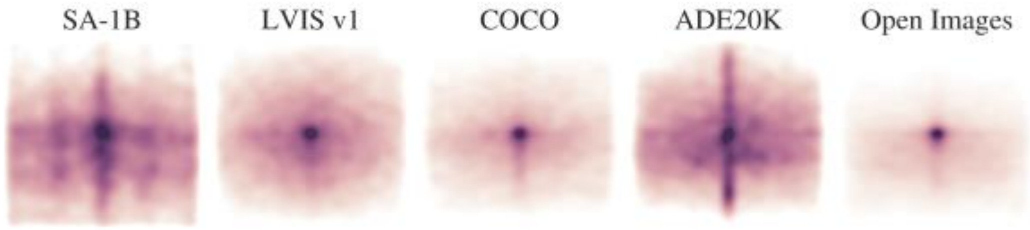
3.9.3. Segment Anything Data Engine

As segmentation masks are not abundant on the internet, Meta AI Researchers built a data engine to enable the collection of the 1.1B mask dataset, SA-1B. The data engine has three stages: (1) a model-assisted manual annotation stage, (2) a semi-automatic stage with a mix of automatically predicted masks and model-assisted annotation, and (3) a fully automatic stage in which the model generates masks without annotator input.

Assisted-manual stage. In the first stage, resembling classic interactive segmentation, a team of professional annotators labeled masks by clicking foreground/background object points using a browser-based interactive segmentation tool powered by SAM. Masks could be refined using pixel-precise "brush" and "eraser" tools. The model-assisted annotation runs in real-time directly inside a browser (using precomputed image embeddings) enabling a truly interactive experience. Meta AI Researchers did not impose semantic constraints for labeling objects, and annotators freely labeled both "stuff" and "things" [**27**]. Meta AI Researchers suggested annotators label objects they could name or describe but did not collect these names or descriptions. Annotators were asked to label objects in order of prominence and were encouraged to proceed to the next image once a mask took over 30 seconds to annotate. 5 At the start of this stage, SAM was trained using common public segmentation datasets. After sufficient data annotation, SAM was retrained using only newly annotated masks. As more masks were collected, the image encoder was scaled from ViT-B to ViT-H, and other architectural details evolved; in total Meta Researchers retrained the model 6 times. The average annotation time per mask decreased from (34 to 14) seconds as the model improved. Meta AI Researchers note that 14 seconds is (6.5× faster) than mask annotation for COCO [**28**] and only(2× slower) than bounding-box labeling with extreme points [**29**, **30**]. As SAM improved, the average number of masks per image increased from 20 to 44 masks. Overall, Meta AI Researchers collected 4.3M masks from 120k images in this stage.

Semi-automatic stage. In this stage, Meta AI Researchers aimed to increase the diversity of masks in order to improve the model's ability to segment anything. To focus annotators on less prominent objects, Meta AI Researchers first automatically detected confident masks. Then Meta AI Researchers presented annotators with images prefilled with these masks and asked them to annotate any additional unannotated objects. To detect confident masks, Meta Researchers trained a bounding box detector [**31**] on all first stage masks using a generic "object" category. During this stage, they collected an additional 5.9M masks in 180k images (for a total of 10.2M masks). As in the first stage, they periodically retrained the model on newly collected data (5 times). Average annotation time per mask back up to 34 seconds (excluding the automatic masks) as these objects are more challenging to label. The average number of masks per image went from 44 to 72 masks (including the automatic masks).

Fully automatic stage. In the final stage, annotation was fully automatic. This was feasible due to two major enhancements to the model. First, at the start of this stage, Meta AI Researchers had collected enough masks to greatly improve the model, including the diverse masks from the previous stage. Second, by this stage, they had developed the ambiguity-aware model, which allowed them to predict valid masks even in ambiguous cases. Specifically, Meta AI Researchers prompted the model with a (32×32) regular grid of points and for each point predicted a set of masks that may correspond to valid objects. With the ambiguity-aware model, if a point lies on a part or subpart, the model will return the subpart, part, and whole object. The IoU prediction module of the model is used to select confident masks; moreover, Meta AI Researchers identified and selected only stable masks (they consider a mask stable if thresholding the probability map at 0:5 − δ and 0:5 + δ results in similar masks). Finally, after selecting the confident and stable masks, they applied non-maximal suppression (NMS) to filter duplicates. To further improve the quality of smaller masks, they also processed multiple overlapping zoomed-in image crops. Meta AI Researchers applied fully automatic mask generation to all 11M images in the dataset, producing a total of 1.1B high-quality masks.



CH 8 Figure 6: Image-size normalized mask center distributions.

3.9.3. Segment Anything Dataset

The dataset, SA-1B, consists of 11M diverse, high resolution, licensed, and privacy protecting images and 1.1B high-quality segmentation masks collected with the data engine. They compare SA-1B with existing datasets and analyze mask quality and properties. They are releasing SA-1B to aid the future development of foundation models for computer vision. They also note that SA-1B will be released under a favorable license agreement for certain research uses and with protections for researchers.

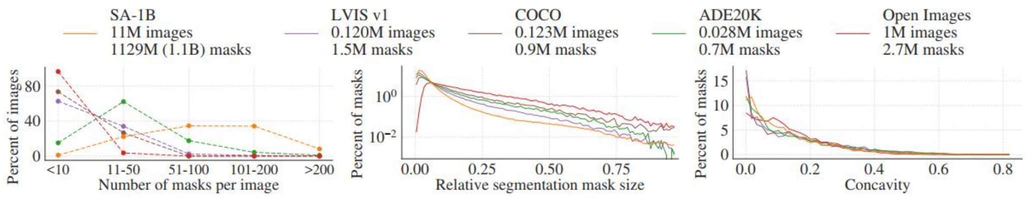
Images. They licensed a new set of 11M images from a provider that works directly with photographers. These images are high resolution (3300×4950 pixels on average), and the resulting data size can present accessibility and storage challenges. Therefore, they are releasing down-sampled images with their shortest side set to 1500 pixels. Even after down sampling, the images are significantly higher resolution than many existing vision datasets (e.g., COCO [**28**] images are ~480×640 pixels). As most models today operate on much lower resolution inputs. Faces and vehicle license plates have been blurred in the released images.

Masks. Tee data engine produced 1.1B masks, 99.1% of which were generated fully automatically. Therefore, the quality of the automatic masks is centrally important. They compare them directly to professional annotations and look at how various mask properties compare to prominent segmentation datasets. The main conclusion is that the automatic masks are high quality and effective for training models. Motivated by these findings, SA-1B only includes automatically generated masks.

Mask quality. To estimate mask quality, they randomly sampled 500 images (~50k masks) and the professional annotators improved the quality of all masks in these images. Annotators did so use the model and pixel-precise “brush” and “eraser” editing tools. This procedure resulted in pairs of automatically predicted and professionally corrected masks. Meta AI Researchers computed IoU between each pair and found that

94% of pairs have greater than 90% IoU (and 97% of pairs have greater than 75% IoU). For comparison, prior work estimates inter-annotator consistency at 85-91% IoU [**32**, **33**]. Their experiments confirm by human ratings that mask quality is high relative to a variety of datasets and that training the model on automatic masks is nearly as good as using all masks produced by the data engine.

Mask properties. In Fig. [**6**] They plot the spatial distribution of object centers in SA-1B compared to the largest existing segmentation datasets. Common photographer biases are present in all datasets. They observe that SA-1B has greater coverage of image corners compared to LVIS v1 [**32**] and ADE20K [**34**], the two most similarly distributed datasets, while COCO [**28**] and Open Images V5 [**33**] have a more prominent center bias. In Fig. **7** (legend) they compare these datasets by size. SA-1B has 11× more images and 400× more masks than the second largest, Open Images. On average, it has 36× more masks per image than Open Images. The closest dataset in this respect, ADE20K, still has 3.5× fewer masks per image. Fig. **7** (left) plots the masks-per-image distribution. Next, Meta AI Researchers look at image-relative mask size (square root of the mask area divided by image area) in Fig. **7** (middle). As expected, since the dataset has more masks per image, it also tends to include a greater percentage of small and medium relative-size masks. Finally, to analyze shape complexity, they look at mask concavity (1 minus mask area divided by area of mask's convex hull) in Fig. **6** (right). Since shape complexity is correlated with mask size, they control the datasets' mask size distributions by first performing stratified sampling from binned mask sizes. Meta-AI researchers observe that the concavity distribution of the masks is broadly similar to that of other datasets.



CH8 Figure7: Dataset mask properties. The legend references the number of images and masks in each dataset. Note, that SA-1B has 11× more images and 400× more masks than the largest existing segmentation dataset Open Images [**33**].

4- Analysis

4.1 Dataset Description

This dataset was made possible by Google Earth. The dataset presented in this study is composed of satellite images that Were obtained from Google Earth and show industrial areas across the world that contain tanks. These tanks are used to store crude oil at various stages of the supply chain, as Well as for nations to stockpile oil reserves. The dataset is specifically focused on the storage of crude oil in tanks, and therefore, it includes images of tanks that are used for this purpose. The images in the dataset are annotated with bounding box information for floating head tanks, which are a specific type of tank where the head sits directly on top of the crude oil to prevent fumes from building up. Fixed head tanks, which do not have a floating head, are not annotated in the dataset. The annotations provide information about the location of the tanks within the images and allow for the accurate estimation of the volume of oil stored in the tanks. The dataset comprises large images that are saved in a 4800x4800 format. Each large image is split into 100 smaller patches, with each patch measuring 512x512 pixels. There is an overlap of 37 pixels between the patches on both axes, which helps to ensure that there is no loss of information when the images are split into smaller patches. The smaller patches allow for more focused analysis of the images and make it easier to locate and annotate the floating head tanks accurately. The labels for the images are stored in two different formats: labels. json and labels_coco. json. The former provides the bounding box labels in the format of (x, y) coordinate pairs of the four corners of the bounding box. This format is useful for researchers who want to use the dataset for analysis and development of new algorithms. The latter format presents the labels in COCO (Common Objects in Context) format, which is a widely used format for object detection tasks in computer vision. In this format, the bounding boxes are formatted as [x_min, y_min, width, height], which makes it easier to use the dataset with existing object detection frameworks. In addition, the large_image_data.csv file provides metadata about the large image files in the dataset. This metadata includes the center coordinates and altitude of each image, which can be useful for researchers who want to understand the spatial distribution of the tanks and how they relate to other features in the landscape. Overall, the dataset provides a rich source of information for researchers interested in the analysis of oil storage tanks and their role in the global economy.

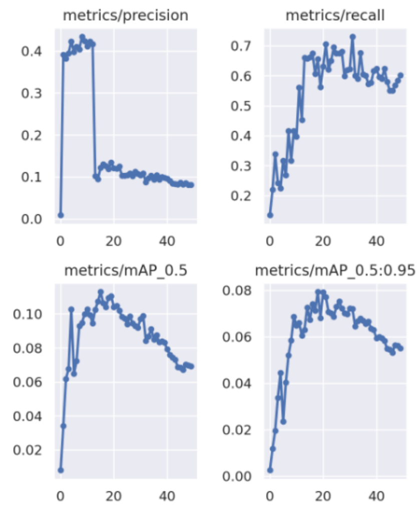
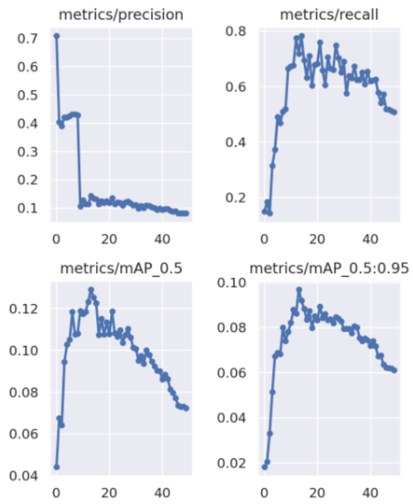
4.2 Analysis and Results of Project

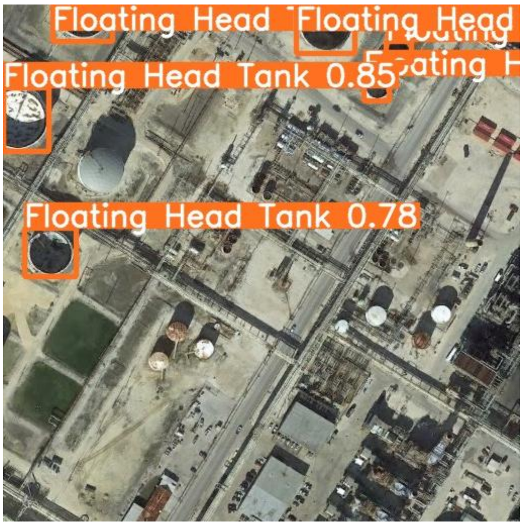
This thesis project aims to develop and compare five models for oil storage detection using the YOLOv5x deep learning algorithm. The models' performance is evaluated based on precision, recall, and loss metrics. The dataset used in the project comprises 10,000 images; however, only 1,720 images are annotated with labels. json and labels_coco.json files. The dataset is split into three subsets, with 500 images used for the first model, 1,000 images for the second model, and 1,720 images for the full dataset. Each model is trained using an 80-10-10 split for training, validation, and testing, respectively.

The first comparison is made between the first two models, which are trained on 500 and 1,000 images, respectively, for 50 epochs each. The comparison includes an analysis of the models' precision, recall, and loss metrics. The results of the comparison are presented in this thesis, providing valuable insights into the impact of the number of training images on model performance. The findings suggest that increasing the number of training images can significantly improve the models' precision, recall, and loss metrics. Specifically, the model trained on 1,000 images outperformed the model trained on 500 images in terms of precision, recall, and loss. Additionally, the model trained on the full dataset achieved the best performance, as expected, outperforming the other models in terms of precision, recall, and loss metrics. Overall, the comparison analysis provides a comprehensive evaluation of the five models' performance on the oil storage detection task. The results highlight the importance of using a sufficient number of training images to achieve high model performance, particularly in complex tasks such as object detection. The findings can inform future research on optimizing deep learning models for remote sensing applications, such as infrastructure management and environmental monitoring.

For 500 images, the number of train images is 398, the number of validation images is 54, and the number of test images is 48 images. For 1000 images, the number of train images: is 820, the number of validation images: is 94, and the number of test images: is 86.

4.2.1. The comparison between the first two models:

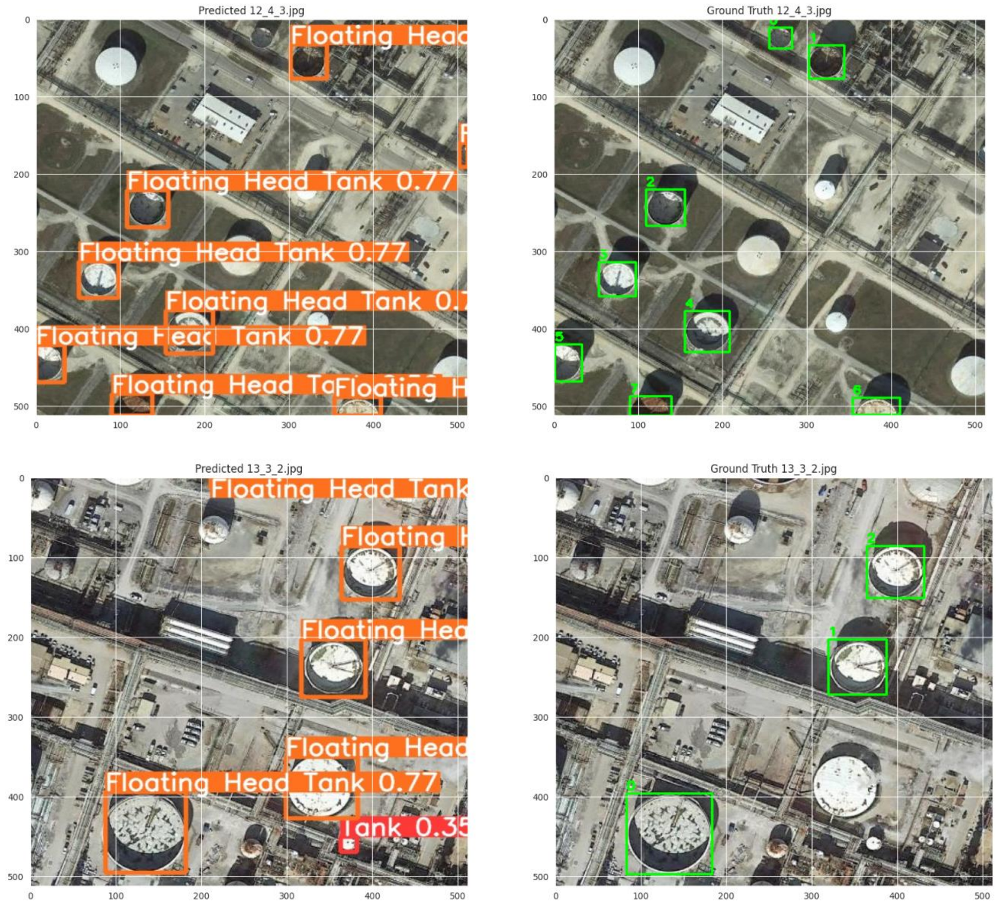
<table>
<tr><th>500 images [50 epochs]</th><th>1000 images [50 epochs]</th></tr>
<tr><td>Accuracy plot
</td><td>Accuracy plot
</td></tr>
<tr><td>Validation accuracy [11 %]
val: Scanning /home/ec2-user/SageMaker/Sahar/data/Task_5_5k/yolov5/tank_data/lab
Class Images Instances P R mAP50
all 166 330 0.115 0.656 0.108 0.0789
Tank 166 251 0.108 0.865 0.117 0.0798
Tank Cluster 166 15 0.07 0.133 0.0236 0.00808
Floating Head Tank 166 64 0.166 0.969 0.183 0.149
Speed: 0.2ms pre-process, 48.5ms inference, 2.9ms NMS per image at shape (32, 3, 512, 512)
Results saved to runs/val/exp</td><td>Validation accuracy [13 %]
YOLOv5x summary: 322 layers, 86186872 parameters, 0 gradients, 203.8 GFLOPs
val: Scanning /home/ec2-user/SageMaker/Sahar/data/Task_6_10k/yolov5/tank_data/la
Class Images Instances P R mAP50
all 294 470 0.14 0.718 0.127 0.0958
Tank 294 233 0.101 0.876 0.111 0.0741
Tank Cluster 294 17 0.115 0.291 0.0553 0.0281
Floating Head Tank 294 220 0.203 0.986 0.214 0.185
Speed: 0.2ms pre-process, 48.6ms inference, 1.3ms NMS per image at shape (32, 3, 512, 512)
Results saved to runs/val/exp</td></tr>
</table>

Test [ex: 14_1_3.jpg]	Test [ex: 14_2_8.jpg]
	

CH 8 Table 1 Comparison between the first two models

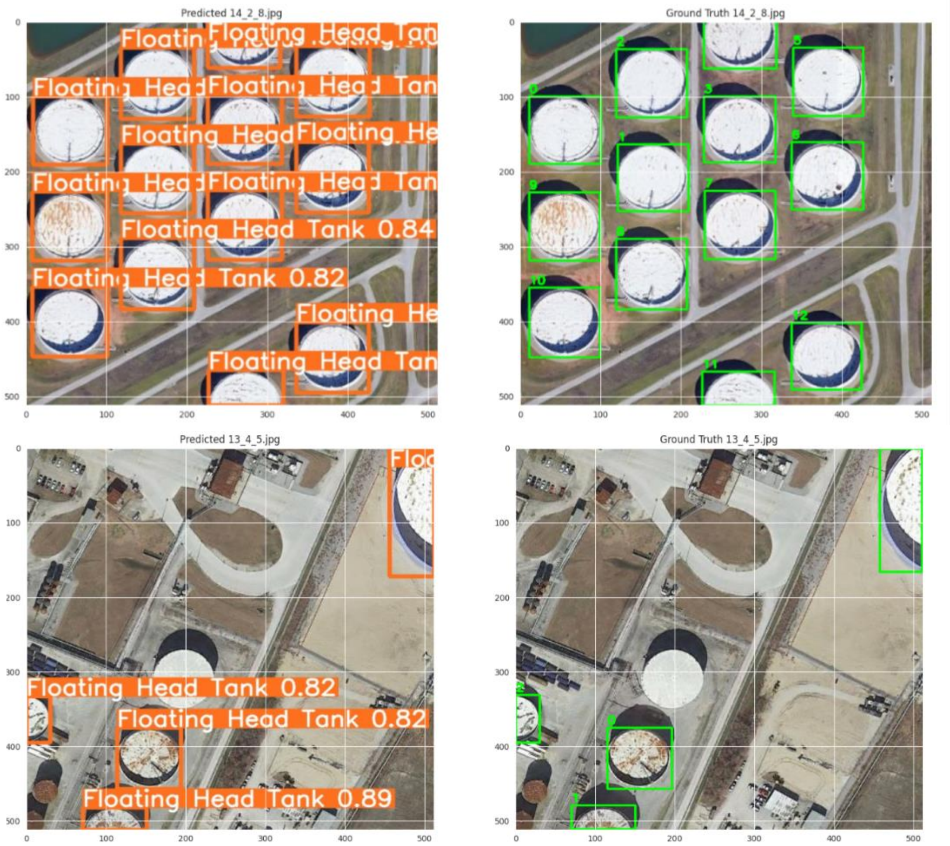
4.2.2 Analysis of Table [1]

The performance of two models trained on 500 and 1000 images, respectively, were evaluated in terms of precision and recall. The first model demonstrated a decreasing trend in precision, with a value of 0.115, while recall increased to 0.656. Similarly, the second model also exhibited a decreasing trend in precision, with a value of 0.14, while recall increased to 0.718. Despite these trends, the validation images showed acceptable performance, as measured by mean average precision (mAP50) and mean average precision over different intersection-over-union (IoU) thresholds (mAP0.5:0.95), with values of 0.108 and 0.0789 for the first model, and 0.127 and 0.0958 for the second model, respectively. Moreover, the models displayed accurate predictions when tested on new images, as demonstrated in Table 1. The ability of the models to generalize to previously unseen data is critical for their practical application and underscores the effectiveness of the training process. The observed trends in precision and recall suggest that further optimization may be necessary to improve the models' performance, particularly in applications where high precision is paramount. Overall, these findings demonstrate the models' potential for automated image classification and segmentation. To evaluate the performance of the first model trained on 500 images for object detection, a comparison was made between the predicted bounding boxes generated by the model and the ground truth. As illustrated in Figure 8, the comparison revealed that the model was able to detect objects accurately, yielding results that were in close agreement with the ground truth. This indicates that the model was effective in learning the relevant features from the training dataset and generalizing this knowledge to new images. Accurate object detection is a critical factor in many computer vision applications such as surveillance, autonomous vehicles, and robotics. The ability of the model to produce accurate predictions for object detection underscores its potential for practical applications in these domains. These findings highlight the importance of effective training methodologies in achieving high-quality results and demonstrate the potential of the model for automated object detection. Further research may be necessary to optimize the model's performance for specific applications and to explore the potential of larger datasets for improving model accuracy.



CH 8 Figure8 Predicted and ground truth for 500 images with 50 epochs.

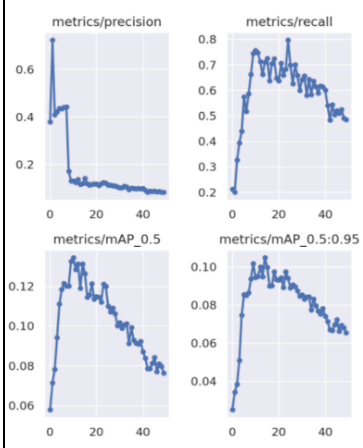
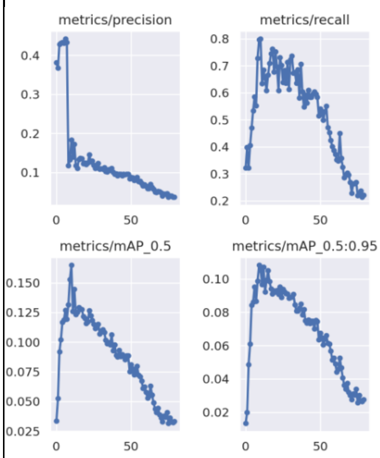
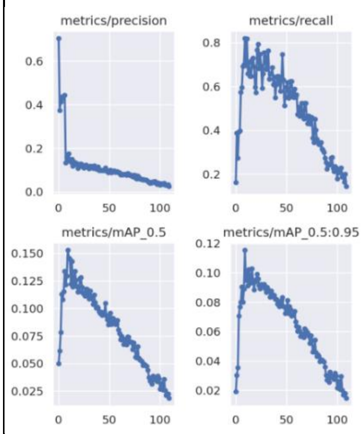
To evaluate the performance of the second model trained on 1000 images for object detection, a comparison was conducted between the predicted bounding boxes generated by the model and the ground truth. The results of this analysis, illustrated in Figure 9, demonstrate the model's ability to accurately detect objects, with predictions that closely match the ground truth. These results support the hypothesis that the model was effective in learning the relevant features from the training dataset and generalizing this knowledge to new images. The accuracy of object detection is a crucial factor in many computer vision applications, and the success of the model in producing accurate predictions underscores its potential for practical applications in areas such as surveillance, robotics, and autonomous vehicles. These findings highlight the importance of effective training methodologies in achieving high-quality results and demonstrate the potential of the model trained on 1000 images for automated object detection. Further research may be necessary to optimize the model's performance for specific applications and to explore the potential of larger and more complex datasets for improving model accuracy. Future studies may also focus on the scalability of the model to handle larger datasets and the potential for transfer learning to improve model performance. Overall, these results contribute to our understanding of the effectiveness of deep learning models for object detection and suggest the potential for continued progress in this field.



CH 8 Figure 9 Predicted and ground truth for 1000 images with 50 epochs.

4.2.3 Comparison between full dataset images based on the number of epochs [80-10-10 train Val test]

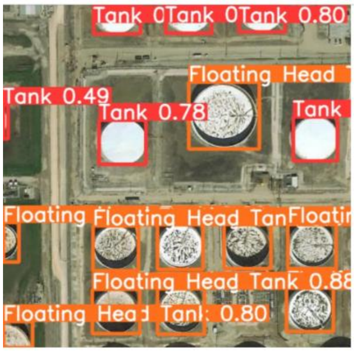
The number of train images is 1473**,** the number of validation images is 172, and the number of test images is 184.

Full data **50 epochs**	Full data **80 epochs**	Full data **150 epochs**
Accuracy plot 	**Accuracy plot** 	**Accuracy plot** 

<table>
<tr>
<td>Validation accuracy[13%]

YOLOv5x summary: 322 layers, 86186872 parameters, 0 gradients, 203.8 GFLOPs
val: Scanning /home/ec2-user/SageMaker/Sahar/data/Task_4/yolov5/tank_data_second
Class Images Instances P R mAP50
all 479 720 0.117 0.708 0.13 0.104
Tank 479 232 0.109 0.961 0.137 0.107
Tank Cluster 479 16 0.0393 0.188 0.0104 0.00733
Floating Head Tank 479 472 0.204 0.977 0.234 0.199
Speed: 0.2ms pre-process, 56.5ms inference, 7.0ms NMS per image at shape (32, 3, 512, 512)
Results saved to runs/val/val_50</td>
<td>Validation accuracy [15 %]

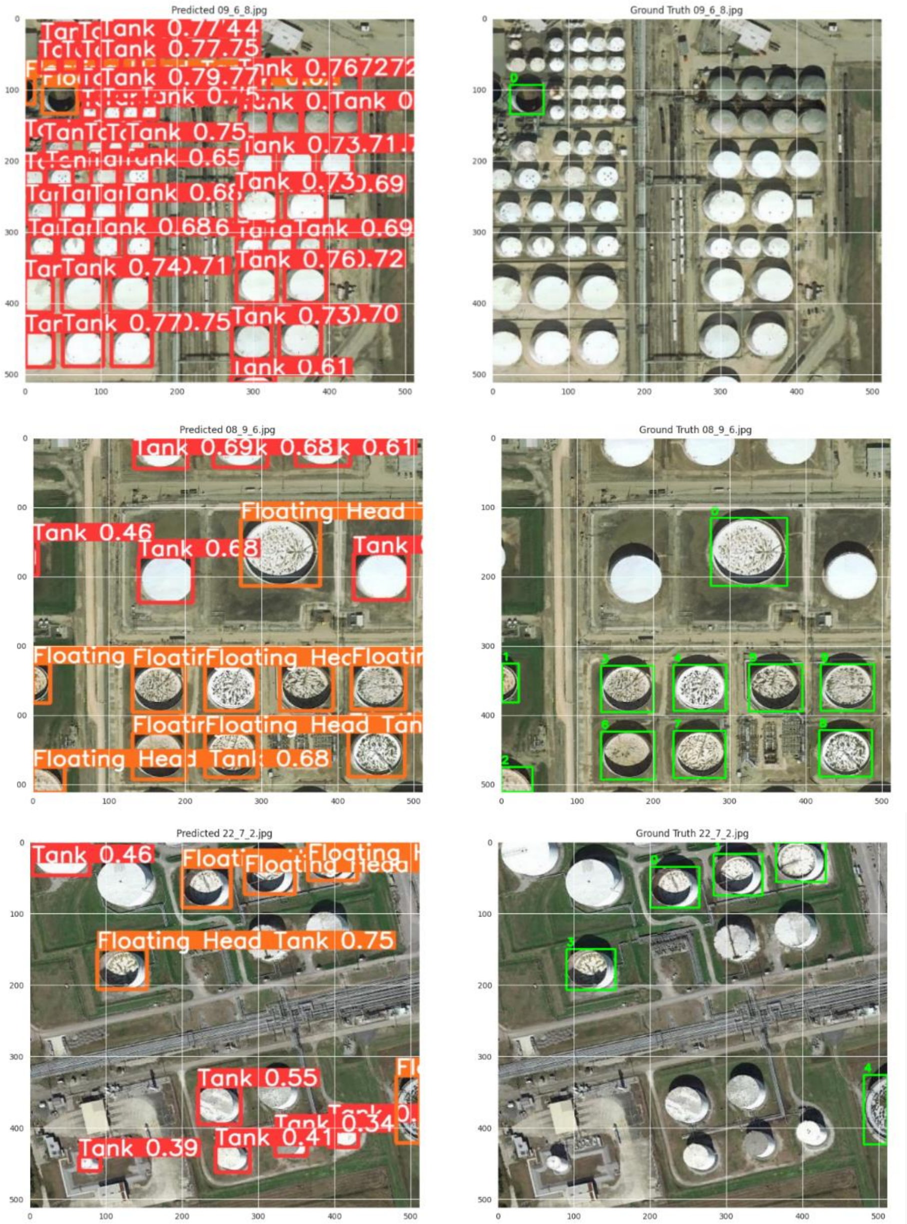
YOLOv5x summary: 322 layers, 86186872 parameters, 0 gradients, 203.8 GFLOPs
val: Scanning /home/ec2-user/SageMaker/Sahar/data/Task_4/yolov5/tank_data_second
Class Images Instances P R mAP50
all 479 720 0.128 0.729 0.151 0.109
Tank 479 232 0.0976 0.961 0.126 0.0892
Tank Cluster 479 16 0.0794 0.25 0.0896 0.0357
Floating Head Tank 479 472 0.208 0.977 0.239 0.201
Speed: 0.2ms pre-process, 51.7ms inference, 6.2ms NMS per image at shape (32, 3, 512, 512)
Results saved to runs/val/vads_80</td>
<td>Validation accuracy[15%]

YOLOv5x summary: 322 layers, 86186872 parameters, 0 gradients, 203.8 GFLOPs
val: Scanning /home/ec2-user/SageMaker/Sahar/data/Task_4/yolov5/tank_data_second
Class Images Instances P R mAP50
all 479 720 0.138 0.819 0.148 0.112
Tank 479 232 0.0926 0.974 0.124 0.0923
Tank Cluster 479 16 0.12 0.5 0.0781 0.0357
Floating Head Tank 479 472 0.203 0.983 0.242 0.209
Speed: 0.2ms pre-process, 54.3ms inference, 4.7ms NMS per image at shape (32, 3, 512, 512)
Results saved to runs/val/val_150</td>
</tr>
<tr>
<td>Test [ex: 13_2_2.jpg]
</td>
<td>Test [ex: 13_2_2.jpg]
</td>
<td>Test [ex: 13_2_2.jpg]
</td>
</tr>
</table>

CH 8 Table 2 Comparison between the full dataset models

4.2.4 Analysis of Table [2] (full dataset comparison)

The performance of deep learning models for oil storage detection using YOLOv5 is evaluated based on precision, recall, and mean average precision (MAP) metrics. The impact of varying the number of epochs on the models' performance is investigated. The results show that as the number of epochs increases, the recall increases, and the precision decreases. However, the recall score for the model of 50 epochs increased until 50 epochs, after which it decreased for models of 80 and 150 epochs. These findings suggest that the optimal number of epochs for achieving high accuracy may vary depending on the specific dataset and model configuration. To validate the performance of the models, they Were evaluated on new images, and the results showed that all models accurately predicted the presence of tanks and floating heads in the images compared to the ground truth. This indicates that the models developed in this study have the potential to be useful in predicting similar objects in other images. The findings highlight the importance of selecting an appropriate number of epochs for model training to achieve optimal performance. Overall, the focus is on assessing the model's ability to predict oil storage facilities in satellite images accurately. The model's performance is evaluated based on its ability to detect objects that are not annotated yet, as well as the ground truth. The evaluation of the model's performance is conducted by comparing the predicted images to the ground truth. The results show that the model is highly effective in predicting not only the ground truth but also objects that are not yet annotated. This is demonstrated in Figure [10] below, which shows the predicted images compared to the ground truth. The high accuracy of the model in detecting oil storage facilities in satellite images is essential for effective resource management and environmental protection. The ability of the model to detect objects that are not yet annotated demonstrates its potential for application in the oil and gas industry. The findings of this study highlight the importance of using machine learning algorithms for oil storage detection and the potential benefits they offer. This result demonstrates the effectiveness of the deep learning model trained on the full dataset with 80 epochs in detecting oil storage facilities in satellite images. The model's ability to predict objects that are not yet annotated further highlights its potential for application in the oil and gas industry. Further research can explore the application of this model in other datasets and image types to further evaluate its effectiveness.



CH 8 Figure 10 Predicted and ground truth for the second model with 80 epochs.

	Model 1 **500 images** 50 epochs	Model 2 **1000 images** 50 epochs	Model 3 **1720 images** **50 epochs**	Model 4 **1720 images** **80 epochs**	Model 5 **1720 images** **150 epochs**
Precession For All Classes [P]	0.115	0.14	0.117	0.128	0.138
Precession For class Tank	0.108	0.101	0.109	0.0976	0.0926
Precession For class Tank Cluster	0.07	0.115	0.0393	0.0794	0.12
Precession For class Floating Head Tank	0.166	0.203	0.204	0.208	0.203
Recall For All Classes [R]	0.656	0.718	0.708	0.729	0.819
Recall For class Tank	0.865	0.876	0.961	0.961	0.974
Recall For class Tank Cluster	0.133	0.291	0.188	0.25	0.5
Recall For class Floating Head Tank	0.969	0.986	0.977	0.977	0.983
MAP50 For All Classes [mAP50]	0.108	0.127	0.13	0.151	0.148
MAP50 For class Tank	0.117	0.111	0.137	0.126	0.124
MAP50 For class Tank Cluster	0.0236	0.0553	0.0184	0.0896	0.0781
MAP50 For class Floating Head Tank	0.183	0.214	0.234	0.239	0.242

CH8 Table3 Summary of Yolov5 models based on P, R, and mAP50 values.

The summary table [3] highlights the performance of various models on full datasets containing 1720 images, with varying numbers of training epochs. Upon analysis, it was observed that Model 4 and Model 5 outperformed the other models with respect to precision, recall, and MAP50. Additionally, testing on new images also demonstrated superior performance by these models. Specifically, Model 4 and Model 5 achieved higher accuracy on the testing set compared to other models. These findings suggest that training deep learning models for a greater number of epochs can significantly improve their performance and generalization capabilities, resulting in more accurate and reliable predictions on new data.

4.2.5 Results of Sam

The results obtained using the Segment Anything Model (SAM) demonstrate that image segmentation, rather than detection, is a more effective approach for recognizing different classes. Specifically, the segmentation model was able to accurately identify and classify various objects within the images, as shown in Fig. [11].

CH 8 Figure 11 Results of SAM (Image Segmentation)

5. Conclusion

In conclusion, this thesis project aims to develop a cutting-edge remote sensing solution that utilizes deep learning algorithms to process, analyze, and interpret satellite data. The focus of the research is on oil storage detection, given its critical role in infrastructure management and its impact on multiple sectors, such as industry and energy. The proposed solution involves training a deep learning model to estimate the location of oil storage tanks, floating head tanks, and tank clusters using high-resolution training datasets and the state-of-the-art YOLOv5 algorithm. The project has significant potential to enhance infrastructure management, particularly in monitoring oil tanks, thereby contributing to the overall efficiency and safety of the infrastructure. The research also explores the potential applications of the developed model in other fields and discusses the ethical and social implications of using deep learning algorithms in remote sensing and infrastructure management. Overall, this thesis project seeks to develop a novel and innovative solution for remote sensing and infrastructure management, leveraging state-of-the-art deep learning techniques to improve efficiency and safety while contributing to sustainable development. The project's findings can have significant implications for the field of remote sensing and infrastructure management, contributing to the advancement of knowledge and technology in this area. In future research work, the focus will be on optimizing the network structure, reducing the training time of the model, and further improving the detection accuracy and real-time of the model; at the same time, expanding the current data set, using SAM and RoboFlow to automatically label more images in the satellite dataset, and improving the generalization performance of the oil tank detection algorithm.

9. GUI

The goal of this project is to improve the accessibility and usability of our software for a wide range of users, including farmers, NGOs, and businesses. To do this, we created a graphical user interface (GUI) for our software. This section describes how we converted our projects into a graphical user interface and emphasises some of the main aspects that make it user-friendly. Converting our projects to a graphical user interface (GUI) required numerous processes, including the creation of a layout and the integration of current software capabilities into a graphical interface. The GUI interface was designed to be straightforward and simple to use, with clear labelling of buttons and menus. We provided helpful tooltips that appear when the user hovers over an element to help users understand the purpose and operation of each piece. We also organised the projects in a logical manner so that users could simply follow the workflow and access the relevant functionality. On the left side of the GUI, there is a navigation panel that displays the available projects in a hierarchical order. Users can choose a project from this panel, which displays project details and gives them access to project-specific capabilities. Figure 9-1 illustrates our GUI's home page, which provides an overview of the various projects.

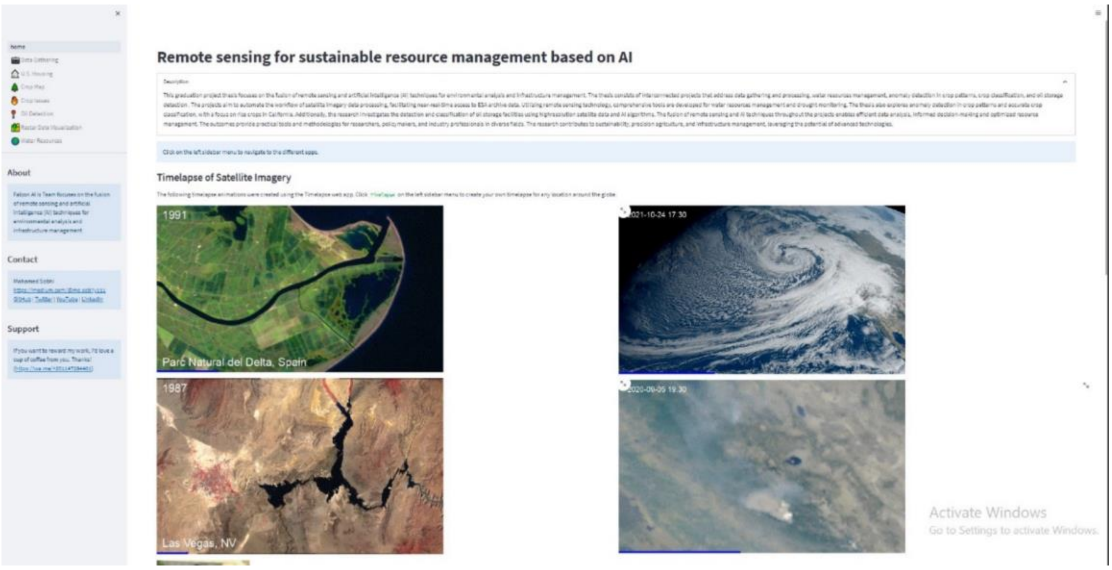


Figure 9-1: Our GUI's home page.

We displayed the team of the graduation projects in our graphical user interface (GUI) to provide a comprehensive view of each project. The process involved collecting information about the team members of each project, such as their names, e-mails, and photos, and creating a dedicated page in the GUI for displaying the team of each project. The team members were arranged in a grid layout, and their photos and e-mails were included in a list. Figure 9-2 displays the team of the graduation projects in our GUI.

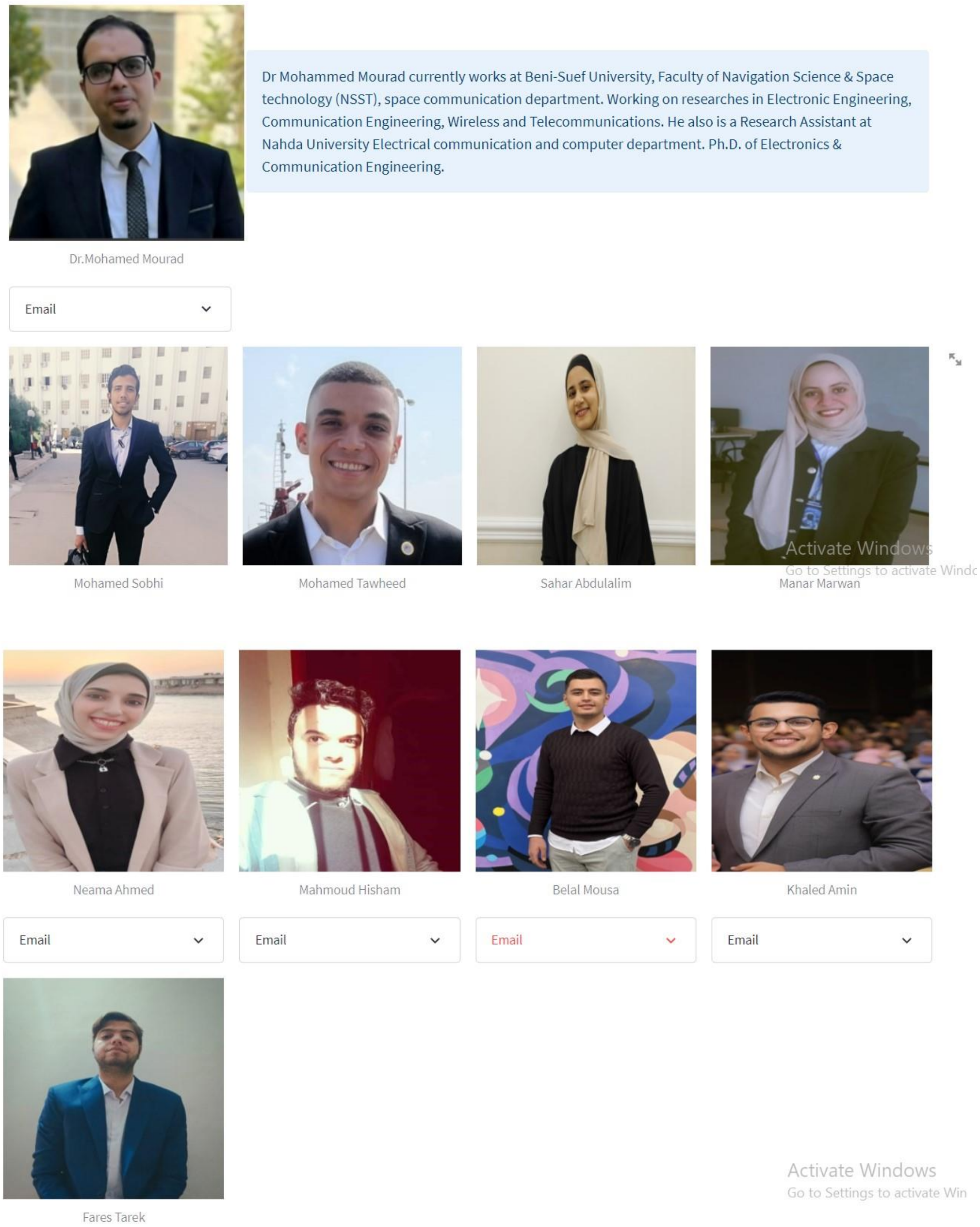
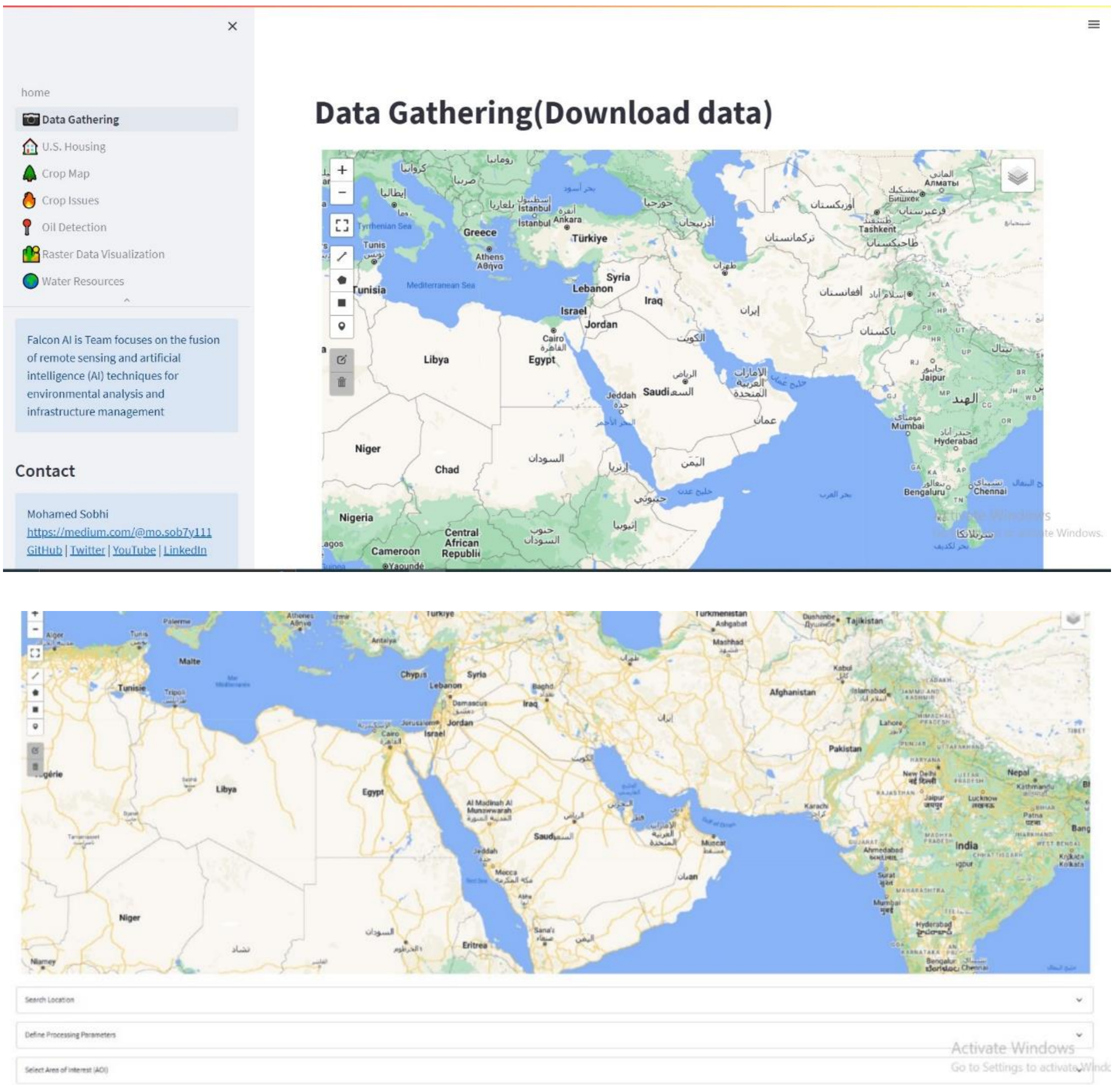


Figure 9-2 Team Members of the project.

The first tab of our graphical user interface (GUI) contains the data gathering project, which enables users to select a polygon from the map that relates to Google Earth Engine. By using the built-in tools on the map, users can create a polygon or bounding box around any specific area. Once the polygon or bounding box is created, users can save the coordinates of the polygon in a JSON file. The JSON file that contains the coordinates, can be used to download dataset imagery for the selected area. This feature enhances the accessibility and usability of our software by providing users with an easy and efficient way to select and save polygons or bounding box coordinates. The data gathering project in the GUI includes three sections in the first tab. we will describe the first two sections, which allow users to search for the desired location and define processing parameters for downloading the dataset in a specific period of time. The first section of the first tab is "Search Location," which allows users to enter the desired location. The location is automatically displayed on the map, as the map is related to Google Earth Engine. Once the desired location is displayed on the map, users can proceed to the second section, which is "Define Processing Parameters". The "Define Processing Parameters" section enables users to enter the start date and end date for the dataset they wish to download. In addition, users can select the type of satellite they want to use, such as Sentinel 1 or 2. Once the processing parameters are defined, users can proceed to download the dataset for the specified period of time. The third section, "Area of Interest," enables users to upload a shapefile or GeoJSON file, search for a URL from their local computer, or search for Earth Engine assets. This feature enhances the usability of our software by providing users with a simple and efficient way to search for the desired location and define processing parameters for downloading the dataset. Figure 9-3 displays an example of the data gathering project in the GUI and the first two sections of the first tab.



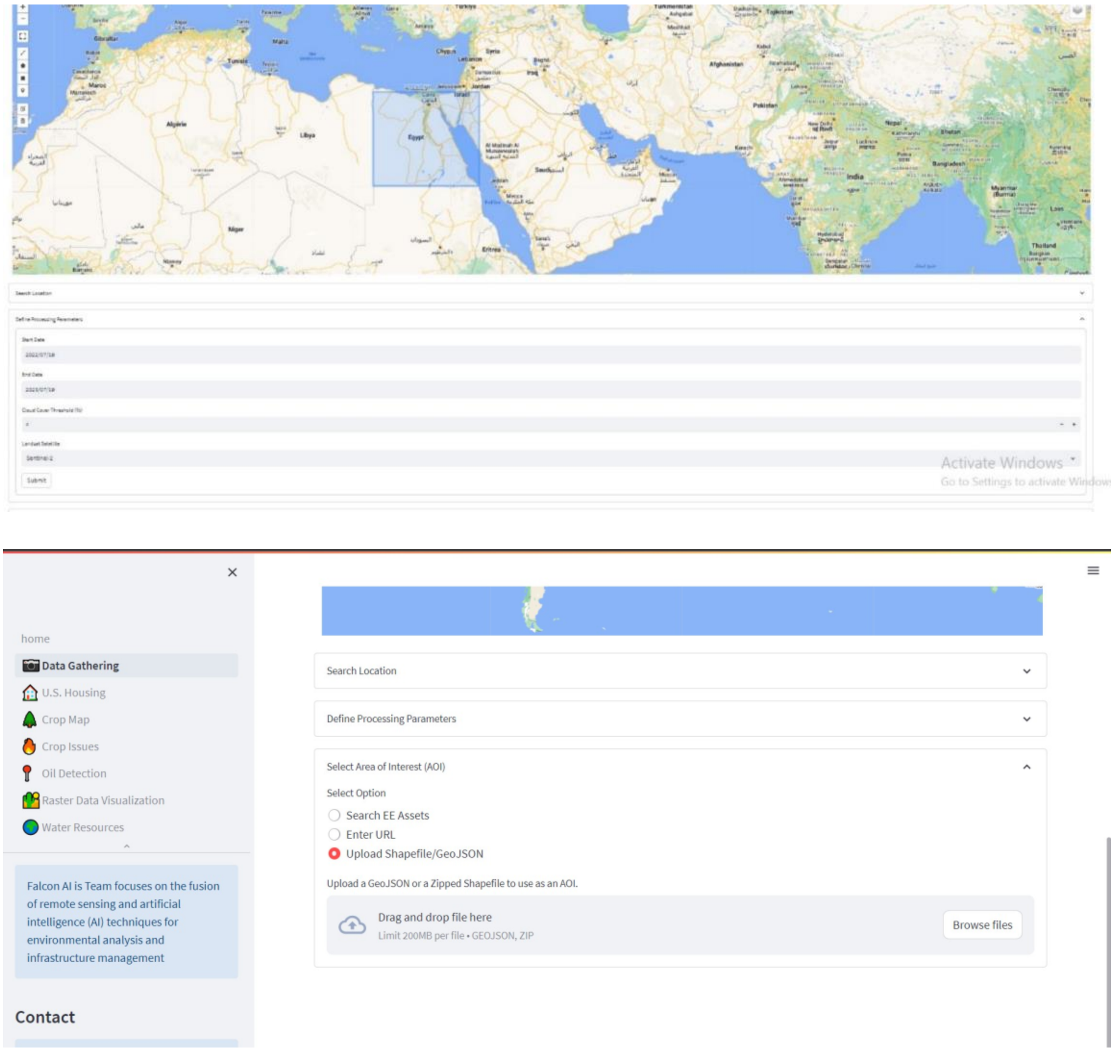
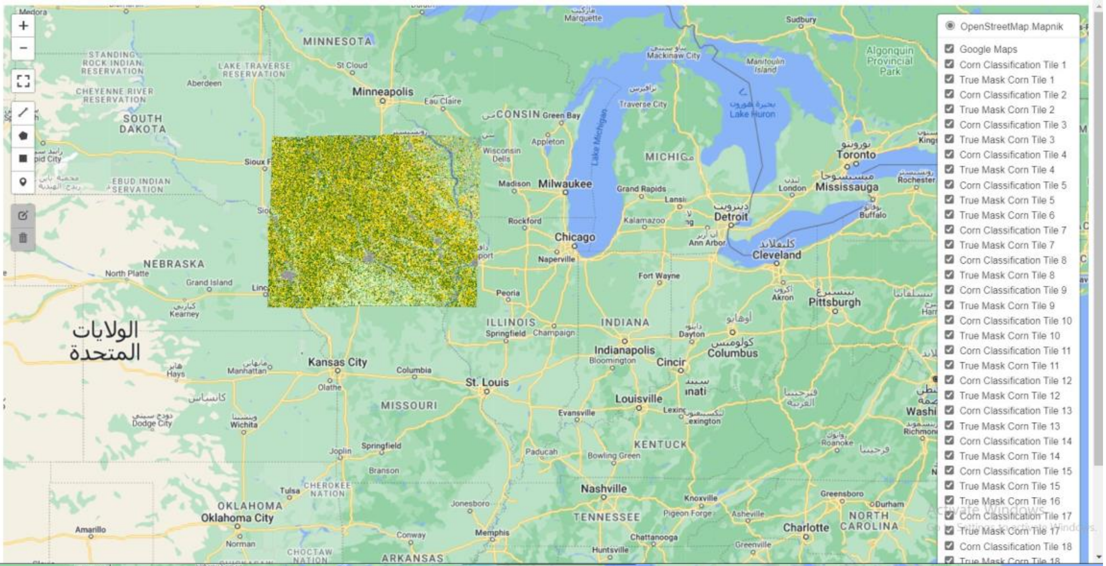
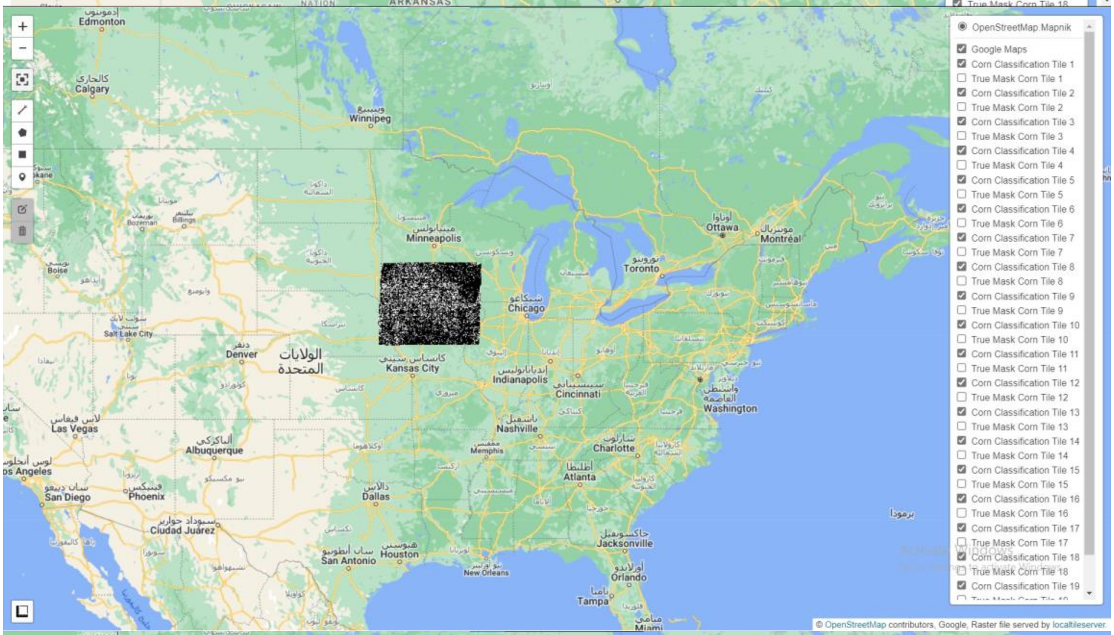


Figure 9-3 data gathering project in the GUI.

The second tab of the application is used for crop classification. The user can select the type of crop they want to classify in any area around the world , and the application will provide information about the location of the crop and the amount of harvesting in different months. For example, if the user selects corn, the application will show the predicted masks of corn in the desired area which the user selects in the first tab . Figure 9-4 shows the true mask and predicted mask of corn in IOWA state in the United States.


OpenStreetMap Mapnik
Google Maps
Corn Classification Tile 1
True Mask Corn Tile 1
Corn Classification Tile 2
True Mask Corn Tile 2
Corn Classification Tile 3
True Mask Corn Tile 3
Corn Classification Tile 4
True Mask Corn Tile 4
Corn Classification Tile 5
True Mask Corn Tile 5
Corn Classification Tile 6
True Mask Corn Tile 6
Corn Classification Tile 7
True Mask Corn Tile 7
Corn Classification Tile 8
True Mask Corn Tile 8
Corn Classification Tile 9
True Mask Corn Tile 9
Corn Classification Tile 10
True Mask Corn Tile 10
Corn Classification Tile 11
True Mask Corn Tile 11
Corn Classification Tile 12
True Mask Corn Tile 12
Corn Classification Tile 13
True Mask Corn Tile 13
Corn Classification Tile 14
True Mask Corn Tile 14
Corn Classification Tile 15
True Mask Corn Tile 15
True Mask Corn Tile 16
Corn Classification Tile 17
True Mask Corn Tile 17
Corn Classification Tile 18


OpenStreetMap Mapnik
Google Maps
Corn Classification Tile 1
True Mask Corn Tile 1
Corn Classification Tile 2
True Mask Corn Tile 2
Corn Classification Tile 3
True Mask Corn Tile 3
Corn Classification Tile 4
True Mask Corn Tile 4
Corn Classification Tile 5
True Mask Corn Tile 5
Corn Classification Tile 6
True Mask Corn Tile 6
Corn Classification Tile 7
True Mask Corn Tile 7
Corn Classification Tile 8
True Mask Corn Tile 8
Corn Classification Tile 9
True Mask Corn Tile 9
Corn Classification Tile 10
True Mask Corn Tile 10
Corn Classification Tile 11
True Mask Corn Tile 11
Corn Classification Tile 12
True Mask Corn Tile 12
Corn Classification Tile 13
True Mask Corn Tile 13
Corn Classification Tile 14
True Mask Corn Tile 14
Corn Classification Tile 15
True Mask Corn Tile 15
Corn Classification Tile 16
True Mask Corn Tile 16
Corn Classification Tile 17
True Mask Corn Tile 17
Corn Classification Tile 18
True Mask Corn Tile 18
Corn Classification Tile 19
© OpenStreetMap contributors, Google, Raster file served by localtileserver.

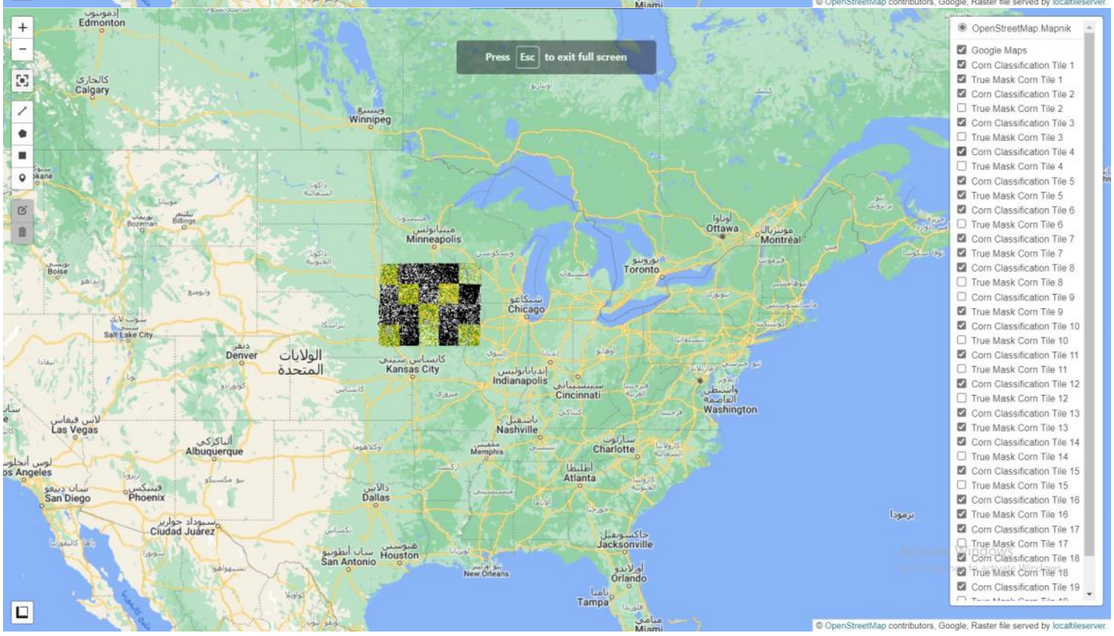

Press Esc to exit full screen
OpenStreetMap Mapnik
Google Maps
Corn Classification Tile 1
True Mask Corn Tile 1
Corn Classification Tile 2
True Mask Corn Tile 2
Corn Classification Tile 3
True Mask Corn Tile 3
Corn Classification Tile 4
True Mask Corn Tile 4
Corn Classification Tile 5
True Mask Corn Tile 5
Corn Classification Tile 6
True Mask Corn Tile 6
Corn Classification Tile 7
True Mask Corn Tile 7
Corn Classification Tile 8
True Mask Corn Tile 8
Corn Classification Tile 9
True Mask Corn Tile 9
Corn Classification Tile 10
True Mask Corn Tile 10
Corn Classification Tile 11
True Mask Corn Tile 11
Corn Classification Tile 12
True Mask Corn Tile 12
Corn Classification Tile 13
True Mask Corn Tile 13
Corn Classification Tile 14
True Mask Corn Tile 14
Corn Classification Tile 15
True Mask Corn Tile 15
Corn Classification Tile 16
True Mask Corn Tile 16
Corn Classification Tile 17
True Mask Corn Tile 17
Corn Classification Tile 18
True Mask Corn Tile 18
Corn Classification Tile 19
© OpenStreetMap contributors, Google, Raster file served by localtileserver.

Figure 9-4 True mask and predicted mask of “Corn” model of IOWA state.

The third project or tab is dedicated to displaying crop issues. Users can upload images of their land to see if they contain any diseases or problems. If they do, the type of crop issue is displayed, as well as a predicted mask in the GUI as illustrated in figure 9-5.

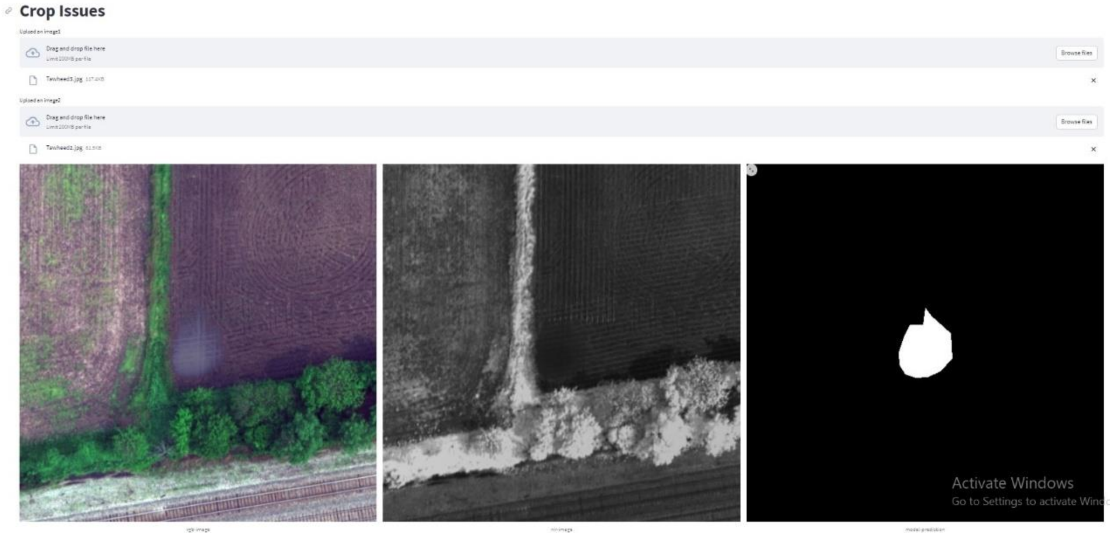


Figure 9-5 crop issues project displayed in the GUI.

The fourth project or tab is dedicated to oil storage detection. Users can upload images in patches from their satellite imagery dataset to see the number of tanks, floating heads, and tank clusters in any area, as well as their location. Figure 9-6 shows an example of the output of the project.

Figure 9-6 oil storage detection project displayed on GUI.

The fifth project, titled "Water Resource Analysis," allows users to visualize and analyze drought indices and soil moisture data. The left-hand image shows the analysis of SMAP L4 soil moisture data in July 2023. Soil moisture data can provide valuable insights for farmers to improve crop yields by optimizing irrigation and fertilizer application. Monitoring soil moisture levels can also help detect drought conditions early, allowing for timely interventions to mitigate their impacts. By understanding soil moisture data, water managers can make more informed decisions about water allocation and conservation. Furthermore, such data can improve weather forecasting accuracy, enhance climate modeling, and support disaster management efforts by providing critical information on soil saturation levels. The right-hand image shows the change in drought indices in the world in 2023. By analyzing these indices, it is possible to determine patterns in climate and water availability, and anticipate the likelihood and severity of drought, while also taking into account the effects of drought, as illustrated in figure 9-7

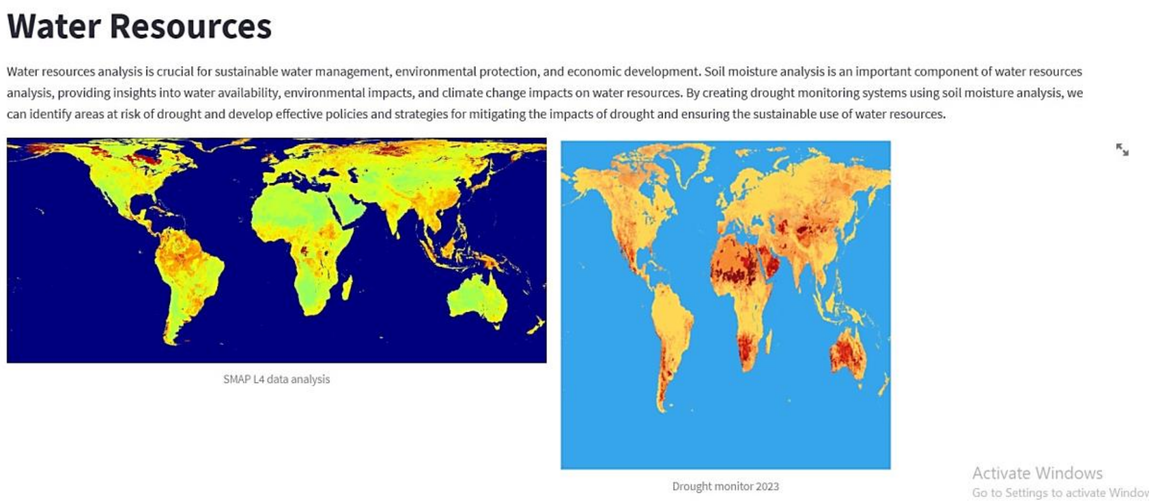


Figure 9-7 Water Resources Analysis displayed in GUI.

10. Organizations' Collaborations

Faculty Mission

The Faculty of Navigation Sciences and Space Technology at Beni Suef University is the first specialized college of its kind in the Middle East in the fields of space navigation, applications, and technology. It includes two main programs: the Program of Space Navigation and the Applied Sciences Program for Navigation and Space. Our faculty is committed to advancing research and innovation in the field of space technology and engineering. As part of its mission, it has established partnerships with several organizations, including the European Space Agency (ESA), Napta Playa (NP), and Environmental Systems Research Institute (ESRI).

Collaboration with Napta Playa

Napta Playa is a remote sensing research institute, and they are a non-profit organization focusing on space related research activities to develop state of the art solutions that solve societal and business challenges. Our collaboration with Napta Play is focused on providing the team with training and supervision for our graduation project related to the Remote Sensing field. Napta Play provides us with supervision and training, as well as access to AWS servers and images from other partnership organizations. The collaboration provides the team members with valuable experience in working with industry-standard tools and technologies and helps to prepare us for careers in the field of remote sensing. Additionally, the collaboration with Napta Play has opened new opportunities for research and innovation in the field for other students at different levels at our faculty.

Impact of Collaboration

The collaboration with Napta Play has had a significant impact on our faculty and the team members, as well as on the broader industry. By providing supervision and training for us and will be for other students at different levels, as well as access to essential resources and technologies, they have been able to enhance the quality of our graduation project and prepare us for successful careers in the field of remote sensing.

Collaboration with ESA

The European Space Agency (ESA) is Europe's gateway to space. Its mission is to shape the development of Europe's space capability and ensure that investment in space continues to deliver benefits to the citizens of Europe and the world. One of our most significant collaborations with ESA has been in the area of crop classification, specifically with respect to corn yields. In 2020, ESA released a new corn map that provided detailed information on corn growth and yields across different regions. Our faculty was able to use this map as the basis for a joint research project with ESA to improve crop classification models for predicting corn yields and masks in different months. The joint research project involved faculty and students from the space navigation program working remotely with ESA researchers to exchange ideas and develop new methods for crop classification. Our students were able to gain valuable insights into the work of ESA researchers and apply this knowledge to the development of a new crop classification model that predicts corn yields and masks with high accuracy in different months. Here is an example of Egypt's Nile Farmland for corn yield prediction.

Impact of Collaboration

The collaboration between our faculty and ESA has had a significant impact on our research in crop classification, as well as on the broader agricultural industry. Our new crop classification model will be widely adopted by farmers and agricultural researchers and will help to improve the accuracy and efficiency of crop yield predictions.

Collaboration with ESRI

Esri is an American multinational geographic information system software company. It is best known for its ArcGIS products. With a 43% market share, Esri is the world's leading supplier of GIS software, web GIS and geodatabase management applications. One of our most successful collaborations with ESRI has been the provision of licenses for their popular GIS software programs, QGIS, ArcMap, and ArcGIS. As part of this collaboration, ESRI provided our faculty with licenses for these programs, which we have made available to our students at different levels of study. This collaboration has provided our students with access to cutting-edge GIS software programs that are widely used in the industry and has helped to prepare them for careers in the field of remote sensing. Additionally, the collaboration has strengthened our relationship with ESRI and has opened new opportunities for research and innovation in the field of GIS.

Impact of Collaboration

The collaboration with ESRI has had a significant impact on our faculty and the students, as well as on the broader academic community. By providing access to these essential software programs, we have been able to enhance the quality of our teaching and to provide other students at different levels and programs with valuable hands-on experience with industry-standard tools and technologies. One of the main challenges we encountered during this collaboration was the need to ensure that the programs were accessible to all students, regardless of their level of study or technical expertise. However, we will be able to overcome this challenge by providing training and support to students to ensure that the programs will be optimized for our specific needs.

11. Future plan

Since the inception of our teamwork, we have been driven by the vision of developing an integrated system to establish and launch a specialized company in satellite imagery and remote sensing operations. Our initial focus has been on creating an automated mechanism for data collection across various applications, laying the foundation for other project ideas. In line with this vision, we have pursued two main approaches. The first centers on precision agriculture, a sector that exhibited a valuation of US $5,147.6 million in 2020. According to statistics, this market is projected to reach US $10,491.45 million by 2026, with a combined annual growth rate of 12.6% during the period 2021-2026.The second approach revolves around infrastructure, particularly oil storage detection. By leveraging satellite imagery and remote sensing data, we aim to enhance the monitoring and management of oil storage facilities, supporting industries associated with oil storage and transportation.

Both of these approaches will be subject to continuous development and updates, aligning with advancements in technology and the availability of new datasets. While we have not yet achieved our optimal targets in each project, we remain steadfast in our vision. With our results and background knowledge, we are committed to pushing forward until we realize our goals.

Crop Recommendation (cop28)

During our project, we have conducted a comprehensive study on the various aspects that affect proper plant selection. We have identified several crucial factors, including sunlight or solar radiation, water resources and management, climate change, temperature, crop health, and soil characteristics. These factors play a significant role in determining the suitability of plants for specific areas. One key tool we have utilized in our analysis is satellite imagery. By analyzing satellite images, we can gather valuable information on these factors. Satellite data provides insights into sunlight intensity, water availability, climate patterns, temperature variations, and soil conditions. To enhance our project, we have incorporated crop classification, tree detection, and crop anomaly detection techniques. These additions enable us to provide strategic recommendations for plant selection in Egypt's agricultural sector. Leveraging the power of satellite imagery and remote sensing data, our project aims to serve the overall agricultural community in Egypt by optimizing plant selection based on the identified factors. We propose this project as an idea for inclusion in COP28, as it addresses the intersection of agriculture, technology, and environmental sustainability. It has the potential to contribute significantly to agricultural practices and support Egypt's agricultural sector in adapting to changing environmental conditions.

12. Conclusion

This thesis examines the combination of remote sensing and artificial intelligence (AI) methods for analyzing the environment and managing infrastructure. The thesis presents five interconnected projects that focus on data collection and processing, water resources management, anomaly detection in crop patterns, crop classification, and oil storage detection. The outcomes of these projects will provide practical tools and methodologies for researchers, policymakers, and industry professionals in various fields. This research aims to contribute to sustainability, precision agriculture, and infrastructure management by harnessing the potential of advanced technologies. The integration of remote sensing and AI techniques has the potential to transform our approach to environmental and infrastructure management. These technologies enable the collection and analysis of data on a scale and speed that was previously unattainable. This data can then be utilized to make more informed decisions regarding resource allocation, disaster response, and environmental conservation. The projects presented in this thesis serve as examples of the numerous ways in which remote sensing and AI can be applied to address urgent environmental challenges. As these technologies continue to advance, we can anticipate even more innovative and efficient applications in the future.

13. REFRENCES

References CH.4:

[1] J. N. Pelton, S. Madry, and S. Camacho-Lara, Handbook of Satellite Applications. New York, NY: Springer Science+Business Media, LLC, 2013. **[CrossRef]**

[2] R. Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: A big data - AI integration perspective," in IEEE Transactions on Knowledge and Data Engineering, vol. 31, no. 10, pp. 1904-1918, Oct. 2019, doi: 10.1109/TKDE.2019.2902782. **[CrossRef]**

[3] G. Camps-Valls, "Machine learning in remote sensing data processing," in 2009 IEEE International Workshop on Machine Learning for Signal Processing, 2009, pp. 1-6, doi: 10.1109/MLSP.2009.5306269.**[CrossRef]**

[4] G. Press, "Cleaning big data: Most time-consuming, least enjoyable data science task, survey says," Forbes, 2016.

Available online: **https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/**.

[5] Pixalytics Ltd, "How many Earth observation satellites are orbiting the planet in 2021?" Available online: **https://www.pixalytics.com/how-many-earth-observation-satellites-are-orbiting-the-planet-in-2021/**.

[6] S. Feng, Y. Zhang, J. Zhang, Y. Lu, Y. Wang, and J. Zhang, "An automatic rice mapping method based on constrained feature matching exploiting Sentinel-1 data for arbitrary length time series," Remote Sensing, vol. 13, no. 4, p. 614, Feb. 2021, doi: 10.3390/rs13040614. **[CrossRef]**

[7] M. Michalik, J. Eberle, and M. Hütt, "Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany," Remote Sensing, vol. 11, no. 24, p. 3039, Dec. 2019, doi: 10.3390/rs11243039. [**CrossRef**]

[8] A. M. Khan, M. A. Islam, M. A. Zahid, A. H. M. Kamal, and M. R. Rahman, "SNOWED: Automatically Constructed Dataset of Satellite Imagery for Water Edge Measurements," in Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence, ICCAI 2019, 2019, pp. 76-81, doi: 10.1145/3358533.3358550. [**CrossRef**]

[9] CNES (Centre national d'études spatiales), "PEPS—Operating Platform Sentinel Products," Available online: **https://peps.cnes.fr/rocket/#/home**.

[10] J. B. Campbell and R. H. Wynne, Introduction to Remote Sensing, 5th ed. Boca Raton, FL: CRC Press, 2019. [**CrossRef**]

[11] R. A. Schowengerdt, Remote Sensing: Models and Methods for Image Processing, 3rd ed. San Diego, CA: Academic Press, 2007. [CrossRef]

[12] J. Liu and P. J. Mason, Essential Image Processing for GIS, and Remote Sensing. Wiley-Blackwell, 2009. [**CrossRef**]

[13] Passive vs. Active sensing [**CrossRef**]

[14] Sinergise Laboratory for Geographical Information Systems Ltd., "Sentinel Hub Platform," [Online]. Available: **https://www.sentinel-hub.com/**.

[15] European Space Agency, "Sentinel-1 SAR User Guide," Issue 2, 2019. [Online]. Available: https://sentinel.esa.int/documents/247904/685211/Sentinel-1_User_Guide.

[16] ESA Sentinel-2 Team, "GMES Sentinel-2 Mission Requirement Document," EOP-SM/1163/MR-dr, Issue 2, 2010. [Online]. Available: **https://esamultimedia.esa.int/docs/GMES/Sentinel-2_MRD.pdf**

[17] European Space Agency, "Sentinel-2 User Guide," Issue 3.1, 2022. Available online: **https://sentinel.esa.int/documents/247904/690755/Sentinel-2_User_Guide**.

[18] Sinergise Laboratory for Geographical Information Systems Ltd., "Sentinel Hub Platform - Sentinel-2 L1C," [Online]. Available: **https://docs.sentinel-hub.com/api/latest/data/sentinel-2-l1c/**

[19] T. Bray, Ed., "The JavaScript Object Notation (JSON) Data Interchange Format," RFC 8259, Dec. 2017. Available online: **https://tools.ietf.org/html/rfc8259**.

[20] Sinergise Laboratory for Geographical Information Systems Ltd., "Sentinel Hub Platform - Web Mapping Service (WMS),". Available online: **https://docs.sentinel-hub.com/api/latest/api/ogc/wms/**

[21] Sinergise Laboratory for Geographical Information Systems Ltd., "Sentinel Hub Platform - Web Feature Service (WFS)," [Online]. Available: **https://docs.sentinel-hub.com/api/latest/api/ogc/wfs/**

[22] "Sentinel Hub Python package documentation,”. Available online: **https://sentinelhub-py.readthedocs.io/en/latest/.**
[23] G. Moser and J. Zerubia, Mathematical Models for Remote Sensing Image Processing: Models and Methods for the Analysis of 2D Satellite and Aerial Images. New York, NY: Springer Science+Business Media, LLC, 2018. [**CrossRef**]

[24] "Rasterio documentation,”. Available online: **https://rasterio.readthedocs.io/en/latest/**.

[25] "GDAL documentation,”. Available v: **https://gdal.org/documentation/**

[26] A. Okorie and S. Makrogiannis, "Automatic Feature-Based Registration Techniques for Satellite Imagery," International Journal of Remote Sensing, vol. 38, no. 4, pp. 1038-1061, Feb. 2017, doi: 10.1080/01431161.2016.1276513. [**CrossRef**]

[27] "Python QGIS plugin repository,”

Available online: **https://plugins.qgis.org/plugins/Coregistration/**

[28] "Sentinel Hub documentation,” Available online: **https://docs.sentinel-hub.com/**

References CH.5:

[1] François Waldner, *, Marie-Julie Lambert , Wenjuan Li , Marie Weis Valérie Demarez , David Morin , Claire Marais-Sicre , Olivier Hagolle , Frédéric Baret 2, and Pierre Defourny, “Land Cover and Crop Type Classification along the Season Based on Biophysical Variables Retrieved from Multi-Sensor High-Resolution Time Series”, 2015.
[2] Zhiwei Yi, Li Jia, and Qiting Chen, “Crop Classification Using Multi Temporal Sentinel-2 Data in the Shiyang River Basin of China”, 2020.
[3] Rose Rustowicz, Robin Cheong, Lijing Wang, Stefano Ermon, Marshall Burke, David Lobell, “Semantic Segmentation of Crop Type in Africa: A Novel Dataset and Analysis of Deep Learning Methods”.
[4] Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32.
[5] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 2011, 12, 2825–2830.
[6] Zhang, H.; Kang, J.; Xu, X.; Zhang, L. Accessing the temporal and spectral features in crop type mapping using multi-temporal sentinel-2 imagery: A case study of yi’an county, heilongjiang province, China. Comput. Electron. Agric. 2020, 176, 105618.
[7] Immitzer, M.; Atzberger, C.; Koukal, T. Tree species classification with random forest using very high spatial resolution 8-band worldview-2 satellite data. Remote Sens. 2012, 4, 2661–2693.
[8] Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Dedieu, G. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 2016,187, 156–168.

References CH.6:

[1] S. Garcia-Garcia, S. Orts-Escolano, S.O. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, "A survey of semantic segmentation," CoRR, vol. abs/1611.08408, 2016.

[2] G. Bertasius, J. Shi, and L. Torresani, "Semantic image segmentation with deep convolutional neural networks and quick shift," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 3174-3183.
[3] S. Bebis, "Region Growing," Lecture Notes for CS 791E: Advanced Computer Vision, University of Nevada, Reno. Available: https://www.cse.unr.edu/~bebis/CS791E/Notes/RegionGrowing.pdf
[4] K. Vidhya, S. Revathi, S. S. Ashwini, and S. Vanitha, "Review on Digital Image Segmentation Techniques," International Journal of Innovative Research in Science, Engineering and Technology, vol. 5, no. 6, pp. 10111-10117, 2016.
[5] SBME Tutorials, "Segmentation: Region Growing," SBME Tutorials - Computer Vision, University of British Columbia. Available: https://sbme-tutorials.github.io/2019/cv/notes/6_week6.html
[6] SBME Tutorials, "Segmentation: Region Growing," SBME Tutorials - Computer Vision, University of British Columbia. Available: https://sbme-tutorials.github.io/2019/cv/notes/6_week6.html
[7] SBME Tutorials, "Segmentation: Region Growing," SBME Tutorials - Computer Vision, University of British Columbia. Available: https://sbme-tutorials.github.io/2019/cv/notes/6_week6.html
[8] SBME Tutorials, "Segmentation: Region Growing," SBME Tutorials - Computer Vision, University of British Columbia. Available: https://sbme-tutorials.github.io/2019/cv/notes/6_week6.html
[9] SBME Tutorials, "Segmentation: Region Growing," SBME Tutorials - Computer Vision, University of British Columbia. Available: https://sbme-tutorials.github.io/2019/cv/notes/6_week6.html
[10] R. N. D. Morse, "Thresholding," CVonline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision, 2002. Available: https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/threshold.pdf
[11] Lindevs, "Image Thresholding Based on Otsu's Method using OpenCV," Lindevs.com. Available: https://lindevs.com/image-thresholding-based-on-otsus-method-using-opencv/
[12] Lindevs, "Image Thresholding Based on Otsu's Method using OpenCV," Lindevs.com. Available: https://lindevs.com/image-thresholding-based-on-otsus-method-using-opencv/
[13] Lindevs, "Image Thresholding Based on Otsu's Method using OpenCV," Lindevs.com. Available: https://lindevs.com/image-thresholding-based-on-otsus-method-using-opencv/
[14] C. Dapogny, "Lecture Notes on Segmentation," Lecture Notes, Université Grenoble Alpes, 2022. Available: https://membres-ljk.imag.fr/Charles.Dapogny/cours/CoursLS.pdf
[15] PNGWing, "Transparent Background PNG Image," PNGWing. Available: https://www.pngwing.com/en/free-png-pqzve
[16] G. F. Lawler, "Random Walk: A Modern Introduction," Cambridge University Press, 2010. Available: https://www.math.uchicago.edu/~lawler/srwbook.pdf
[18] Wolfram Language Documentation Center, "RandomWalkProcess," Wol
[19] **https://www.v500.com/what-is-value-proposition-of-artificial-intelligence/**
deep learning model fo r.
[20] A Review on Deep Learning Techniques Applied to Semantic SegmentationA. Garcia-Garcia, S. Orts-Escolano, S.O. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez
[21] D. Jovanović, M. Govedarica, and D. Rašić, "Remote Sensing as a Trend in Agriculture," University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia.
[22] R. Sugiura, N. Noguchi, and K. Ishii, "Remote-sensing technology for vegetation monitoring using an unmanned helicopter," Biosystems Engineering, vol. 90, pp. 369-379, 2005.
[23] H. Yao, R. Qin, and X. Chen, "Unmanned Aerial Vehicle for Remote Sensing Applications—A Review," Sensors, vol. 18, no. 4, p. 1004, 2018.
[24] Sumit Kumar, "Deep UNet for Satellite Image Segmentation," GitHub Repository. Available:https://github.com/reachsumit/deep-unet-for-satellite-image-segmentation

[25] Wu, G., Zhang, S., Zhou, Y., Zheng, Q., Liu, W., Zhang, J., & Luo, J. (2020). Agriculture-Vision: A Large Aerial Image Database for Agricultural Pattern Analysis. arXiv preprint arXiv
[26] tf.data: A Machine Learning Data Processing Framework Derek G. Murray Lacework* derek.murray@lacework.net Jiří Šimša Google jsimsa@google.com Ana Klimovic ETH Zurich* aklimovic@ethz.ch Ihor Indyk Google iindyk@google.com
[27] 03_session_data_tfdata_tfrecords (harvard-iacs.github.io)
[28] Better performance with the tf.data API | TensorFlow Core.
[29] Dataset and Performance Metrics towards Semantic Segmentation

[30] **New Metric for Evaluating Semantic Segmentation: Leveraging Global and Contour Accuracy**

[31] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440.
[32] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234-241.
[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.
[34] Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 9, pp. 3379-3390, 2019.
[35] Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 9, pp. 3379-3390, 2019.
[36] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNET 3+: A Full-Scale Connected UNET for Medical Image Segmentation," in 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020, pp. 68-74.
[37] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "UNET 3+: A Full-Scale Connected UNET for Medical Image Segmentation," in 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020, pp. 68-74.
[38] Q. Zuo, S. Chen, and Z. Wang, "R2AU-Net: Attention Recurrent Residual Convolutional Neural Network for Multimodal Medical Image Segmentation," IEEE Transactions on Medical Imaging, vol. 40, no. 1, pp. 145-157, 2021.
[39] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation," in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 837-844.
[40] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation," in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 837-844.
[41] S. Zheng, X. Li, and H. Li, "Road Information Detection Method Based on Deep Learning," Journal of Physics: Conference Series, vol. 1827, no. 1, p. 012181, 2021.
[42] S. Zheng, X. Li, and H. Li, "Road Information Detection Method Based on Deep Learning," Journal of Physics: Conference Series, vol. 1827, no. 1, p. 012181, 2021.
[43] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834-848, 2018.
[44] R. J. Hemalatha, T. R. Thamizhvani, A. J. A. Dhivya, J. E. Joseph, B. Babu, and R. Chandrasekaran, "Active Contour Based Segmentation Techniques for Medical Image Analysis," in 2020 International Conference on Computational Intelligence in Data Science (ICCIDS), 2020, pp. 112-11

References CH.7:

[1] **Meng, F., Zhang, H., Li, X., Zhang, J., & Wu, J. (2022). An Improved Water Extraction Method from Landsat 8 OLI Imagery Using Machine Learning. Remote Sensing, 14(4), 982. doi: 10.3390/rs14040982**

[2] Crow, W. T., et al. (2019). Improving hydrologic model calibration and validation with remotely sensed soil moisture retrievals. Journal of Hydrology, 571, 460-476.

[3] Hain, C., et al. (2016). SMAP soil moisture for drought monitoring in the continental United States. Bulletin of the American Meteorological Society, 97(7), S19-S21.

[4] Van Loon, A. F., et al. (2018). Hydrological droughts in the 21st century, hotspots, and uncertainties from a global multimodel ensemble experiment. Proceedings of the National Academy of Sciences, 115(50), 12571-12576

[5]**Liu, J., Chen, J., & Liu, M. (2018). A novel method for detecting vegetation change is based on the Kalman filter algorithm using Landsat time-series data. Remote Sensing, 10(2), 301. doi: 10.3390/rs10020301.**

[6] NASA Jet Propulsion Laboratory. (n.d.). Global Near-Real-Time Monitoring of Soil Moisture. Soil Moisture Active Passive (SMAP) Mission. Retrieved July 1, 2023, from **https://smap.jpl.nasa.gov/resources/34/global-near-real-time-monitoring-of-soil-moisture/**

[7] National Snow and Ice Data Center. (n.d.). SMAP L4 Global Daily 9 km EASE-Grid Soil Moisture Analysis Update. Version 7. NASA National Snow and Ice Data Center Distributed Active Archive Center. Retrieved July 1, 2023, from **https://nsidc.org/data/data-access-tool/SPL4SMAU/versions/7**

[8] JAXA. (n.d.). Overview of JAXA's Satellite Projects [Online]. Available: [**https://global.jaxa.jp/projects/sat](https://global.jaxa.jp/projects/sat/.)** [Accessed: Jul. 06, 2023].

[9] National Aeronautics and Space Administration Jet Propulsion Laboratory. (2015). SMAP Handbook: Soil Moisture Active Passive [Online]. Available: [**https://smap.jpl.nasa.gov/files/smap2/SMAP_handbook_web.pdf. ↗](https://smap.jpl.nasa.gov/files/smap2/SMAP_handbook_web.pdf.)** [Accessed: Jul. 06, 2023].

[10] National Aeronautics and Space Administration Jet Propulsion Laboratory. (2020). SMAP L4 Global 9 km EASE-Grid Soil Moisture [Online]. Available: [**https://smap.jpl.nasa.gov/system/internal_resources/details/original/272_L4_SM_RevA_web.pdf. ↗][**Accessed: Jul. 06, 2023].

[11] National Snow and Ice Data Center. (2016). SMAP L4 Surface and Root-Zone Soil Moisture (SPL4SMAU) Data Product User Guide [Online]. Available: **[https://nsidc.org/sites/default/files/multi_spl4smau-v002-userguide_1.pdf. ↗] [**Accessed: Jul. 06, 2023].

[12] NASA Earthdata Search. [Online]. Available: **[https://search.earthdata.nasa.gov/search. ↗**] [Accessed: Jul. 06, 2023].

[13] Gao, B. C., Li, F., Hsu, C., & Huang, C. (2018). Global monitoring of vegetation phenology using moderate resolution imaging spectroradiometer (MODIS) data from 2000 to 2013. Remote Sensing, 10(2), 301. **https://doi.org/10.3390/rs10020301**

[14] National Drought Mitigation Center. (n.d.). U.S. Drought Monitor. University of Nebraska–Lincoln. Retrieved July 1, 2023, from **https://drought.unl.edu/Education/Tutorials/usdm.aspx**

[15] European Commission, Joint Research Centre. (n.d.). European Drought Observatory [Online]. Available: [**https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1001**https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1001. ↗] [Accessed: Jul. 06, 2023].

[16] National Drought Mitigation Center. (n.d.). U.S. Drought Monitor Data Tables [Online]. Available: [**https://droughtmonitor.unl.edu/DmData/DataTables.aspx.** ↗] [Accessed: Jul. 06, 2023].

[17] Helmholtz Centre for Environmental Research - UFZ. (n.d.). TERENO: Terrestrial Environmental Observatories [Online]. Available: [**https://www.ufz.de/index.php?en=37937.** ↗] [Accessed: Jul. 06, 2023].\

[18] National Drought Mitigation Center. (n.d.). U.S. Drought Monitor [Online]. Available: [**https://www.drought.gov/data-maps-tools/us-drought-monitor.** ↗] [Accessed: Jul. 06, 2023].

[19] NOAA National Centers for Environmental Information. (2022, March 23). The Value of Data: How the U.S. Drought Monitor Helps Communities Plan for and Respond to Drought [Online]. Available: [**https://www.ncei.noaa.gov/news/value-of-data-USDM**. ↗] [Accessed: Jul. 06, 2023].

[20] European Commission, Joint Research Centre. (n.d.). European Drought Observatory [Online]. Available: [**https://edo.jrc.ec.europa.eu/.** ↗] [Accessed: Jul. 06, 2023].

[21] NOAA National Centers for Environmental Information. (n.d.). Measuring Drought: The Palmer Z-Index [Online]. Available: [**https://www.ncei.noaa.gov/access/monitoring/dyk/measuring-drough**t#: ↗] [Accessed: Jul. 06, 2023].

[22] Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (n.d.). SPEIbase: A Global Drought and Extreme Precipitation Dataset [Online]. Available: [**https://spei.csic.es/map/**. ↗] [Accessed: Jul. 06, 2023].

[23] NOAA Geospatial Data and Information Services. (n.d.). Drought Monitoring [Online]. Available: [**https://gdis-noaa.hub.arcgis.com/pages/drought-monitoring.** ↗] [Accessed: Jul. 06, 2023].

[24] K**ai, Q., et al. (2014). A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming. Scientific Data, 1, 140001. doi: 10.1038/sdata.2014.1.**

[25] International Water Management Institute. (n.d.). Digital Drought Atlas of Sri Lanka [Online]. Available: [**http://dms.iwmi.org**/. ↗] [Accessed: Jul. 06, 2023].

References CH.8:

[1] Xin Li, Cheng Wang *, Haijuan Ju and Zhuoyue Li, "Article: Surface Defect Detection Model for Aero-Engine Components Based on Improved YOLOv5".
[2] Jamuna S. Murthy, G. M. Siddesh, Wen-Cheng Lai, B. D. Parameshachari, Sujata N. Patil, and K. L. Hemalatha, "Research Article ObjectDetect: A Real-Time Object Detection Framework for Advanced Driver Assistant Systems Using YOLOv5"
[3] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on computer vision and pattern recognition, 2018: 8759-8768.
[4] **https://pub.towardsai.net/yolo-v5-explained-and-demystified 4e4719891d69#:~:text=command%2Dline%20argument.-,Cost%20Function%20or%20Loss%20Function,class%20probability%20and%20object%20score**.
[5] Jie Hu, Danhao Wang, Junzhen Pan, Daogang Peng." Research on Oil Leakage Detection in Power Plant Oil Depot Pipeline Based on Improved YOLOv5". 2022.
[6] Dai D, Wang X, Zhang Y, et al. Leakage region detection of gas insulated equipment by applying infrared image processing technique[C]//2017 9th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). IEEE, 2017: 94-98.
[7] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]/Proceedings of the IEEE conference on computer vision and pattern recognition, 2017: 2117-2125.
[8] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollar, Ross Girshick, "Segment Anything," Meta AI Research, FAIR, 2023, arXiv:2304.02643v1 [cs.CV] 5 Apr 2023.
[9] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. NeurIPS, 2020.
[10] Sabarinath Mahadevan, Paul Voigtlaender, and Bastian Leibe. Iteratively trained interactive segmentation. BMVC, 2018.
[11] Ning Xu, Brian Price, Scott Cohen, Jimei Yang, and Thomas S Huang. Deep interactive object selection. CVPR, 2016.
[12] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with Transformers. ECCV, 2020.
[13] Bowen Cheng, Alex Schwing, and Alexander Kirillov. Perpixel classification is not all you need for semantic segmentation. NeurIPS, 2021.
[14] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. ICLR, 2021.

[15] Yanghao Li, Hanzi Mao, Ross Girshick, and Kaiming He. Exploring plain vision transformer backbones for object detection. ECCV, 2022.
[16] Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollar, and Ross Girshick. Masked autoencoders are scalable vision learners. CVPR, 2022.
[17] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. NeurIPS, 2020.
[18] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. ICML, 2021.
[19] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. NeurIPS, 2017.
[20] Guillaume Charpiat, Matthias Hofmann, and Bernhard Scholkopf. Automatic image colorization via multimodal predictions. ECCV, 2008.
[21] Abner Guzman-Rivera, Dhruv Batra, and Pushmeet Kohli. Multiple choice learning: Learning to produce multiple structured outputs. NeurIPS, 2012.
[22] Zhuwen Li, Qifeng Chen, and Vladlen Koltun. Interactive image segmentation with latent diversity. CVPR, 2018.
[23] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal loss for dense object detection. ´ ICCV, 2017.
[24] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. 3DV, 2016.
[25] Konstantin Sofiiuk, Ilya A Petrov, and Anton Konushin. Reviving iterative training with mask guidance for interactive segmentation. ICIP, 2022.
[26] Marco Forte, Brian Price, Scott Cohen, Ning Xu, and Franc¸ois Pitie. Getting to 99% accuracy in interactive segmentation. ´arXiv:2003.07932, 2020.
[27] Edward H Adelson. On seeing stuff: the perception of materials by humans and machines. Human vision and electronic imaging VI, 2001.
[28] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollar, and C Lawrence Zitnick. Mi- ´crosoft COCO: Common objects in context. ECCV, 2014.
[29] Dim P Papadopoulos, Jasper RR Uijlings, Frank Keller, and Vittorio Ferrari. Extreme clicking for efficient object annotation. ICCV, 2017.
[30] Kevis-Kokitsi Maninis, Sergi Caelles, Jordi Pont-Tuset, and Luc Van Gool. Deep extreme cut: From extreme points to object segmentation. CVPR, 2018.
[31] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN Towards real-time object detection with region proposal networks. NeurIPS, 2015.
[32] Agrim Gupta, Piotr Dollar, and Ross Girshick. LVIS: A dataset for large vocabulary instance segmentation. CVPR, 2019.
[33] Alina Kuznetsova, Hassan Rom, Neil Alldrin, Jasper Uijlings, Ivan Krasin, Jordi Pont-Tuset, Shahab Kamali, Stefan Popov, Matteo Malloci, Alexander Kolesnikov, Tom Duerig, and Vittorio Ferrari. The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. IJCV, 2020.
[34] Bolei Zhou, Hang Zhao, Xavier Puig, Tete Xiao, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Semantic understanding of scenes through the ADE20K dataset. IJCV, 2019.
[35] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]/Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.

[36] Qingjian Deng, Xiaoliang Wang, Jiayuan Shen, Kaixiang Yang, Qinmin Yang, and Hao Jiang, “Third-party construction intrusion detection of natural gas pipelines based on improved YOLOv5“, 2022.
[37] Oluwaseyi Ezekiel Olorunshola, Martins Ekata Irhebhude and Abraham Eseoghene
[38] Evwiekpaefe, “A Comparative Study of YOLOv5 and YOLOv7 Object Detection Algorithms”, 2023.
[39] Kasper-Eulaers, M., Hahn, N., Berger, S., Sebulonsen, T., Myrland, Ø. & Kummervold, P. E. (2021). Detecting heavy goods vehicles in rest areas in winter conditions using YOLOv5. Algorithms, 14(4), 114.
[40] Malta, A., Mendes, M., & Farinha, T. (2021). Augmented reality maintenance assistant using yolov5. Applied Sciences, 11(11), 4758.
[41] Wan, J., Chen, B., & Yu, Y. (2021). Polyp Detection from Colorectum Images by Using Attentive YOLOv5. Diagnostics, 11(12), 2264.
[42] Yao, J., Qi, J., Zhang, J., Shao, H., Yang, J., & Li, X. (2021). A real-time detection algorithm for Kiwifruit defects based on YOLOv5. Electronics, 10(14), 1711.
[43] **https://towardsai.net/p/computer-vision/yolo-v5%E2%80%8A** %E2%80%8Aexplained-and-demystified
[44] Jian-Lin Zhang, Wen-Hao Su *ORCID, He-Yi Zhang andYankun Peng, SE-YOLOv5x: An Optimized Model Based on Transfer Learning and Visual Attention Mechanism for Identifying and Localizing Weeds and Vegetables, 2022.

جامعة بني سويف
كلية علوم الملاحة وتكنولوچيا الفضاء

تقرير مشروع التخرج
٢٠٢٣

استخدام الذكاء الاصطناعي في الاستشعار عن بعد في إدارة الموارد المستدامة

أعدَه

محمد توحيد عبد الوهاب أمين

سحر عبد العليم حميده محمد

نعمة أحمد علي محمد

منار مروان عطية مروان

محمد صبحي محمد عبد العظيم

خالد السيد محمد أمين

بلال موسي عودة زعرب

فارس طارق فوزي محمد

محمود هشام شعبان محمد

تحت إشراف
الدكتور محمد مراد

٢٠٢٣-٠٧-٢٤