

MONITORING DEFORESTATION IN JORDAN USING SEMANTIC SEGMENTATION AND DEEP LEARNING

**MONITORING DEFORESTATION IN JORDAN USING SEMANTIC
SEGMENTATION AND DEEP LEARNING**

By

Lujain Hamed Mohammad Al-Smadi

Advisor

Dr. Ahmad Gazi Alzu’bi

Thesis submitted in partial fulfillment of the requirements for the degree of
M.Sc. in Computer Science

At

The Faculty of Graduate Studies
Jordan University of Science and Technology

August, 2021

MONITORING DEFORESTATION IN JORDAN USING SEMANTIC SEGMENTATION AND DEEP LEARNING

By

Lujain Hamed Mohammad Al-Smadi

Signature of Author

.....

Committee Members

Signature and Date

Dr. Ahmad Gazi Alzu'bi (Chairman)

.....

Dr. Qanita Bani Baker (Member)

.....

Prof. Faisal Alkhateeb (External Examiner)

.....

August, 2021

تفويض

نحن الموقعين أدناه، نتعهد بمنح جامعة العلوم والتكنولوجيا الاردنية حرية التصرف في نشر محتوى الرسالة الجامعية، بحيث تعود حقوق الملكية الفكرية لرسالة الماجستير إلى الجامعة وفق القوانين والأنظمة والتعليمات المتعلقة بالملكية الفكرية وبراءة الاختراع.

الطالب	المشرف الرئيس
لوجين حامد الصمادي	د. أحمد غازي الزعبي
الرقم الجامعي والتوقيع 121868 	التوقيع والتاريخ

DEDICATION

I dedicate the work done in this thesis

To the one who encouraged me to persevere all my life, to the sea of love and tenderness and the pulse that dwells in my heart, to my beacon that lights my path, to the one who gave me and still gives me without limits, to whom I am proud with,

*Dear Father Hamed AL-Smadi
to the paradise of the world and the apple of my eyes*

Dear Mother

*Your prayers were the reason for what I am now
To my unique and wonderful supervisor
To My Grandfather's Soul*

To those whose love runs through my veins, brothers, sisters and dear uncle

Adnan

Without whom none of my success would be possible

*Lujain Hamed Mohammad Al-Smadi
Jordan, 2021*

ACKNOWLEDGMENT

In the Name of Allah, the Most Merciful, the Most Compassionate. All praise be to Allah, the Lord of the worlds. Prayers and peace be upon Mohamed His servant and messenger. First and foremost, I must acknowledge my limitless and endless thanks to Allah, the Ever-Magnificent; the Ever-Thankful, for His help and bless. I am totally sure that this work would have never become truth, without His guidance.

First and foremost, I have to thank my research supervisor, Dr Ahmad Alzu'bi, Without his assistance and dedicated involvement in every step throughout the process, this thesis would have never been accomplished. I would like to thank you very much for your support and understanding over these past months.

I am pleased to extend my heartfelt thanks to my father and mother for their continuous support, prayers and patience for all that I have been through. I thank my little sister, Asal, for helping me and staying up late with me during the dataset collection period, and my brothers and sisters Montaser, Mohammad, Omar, Malak and Leen for being by my side. Thanks to my dear uncle Adnan for his continuous moral support. I will not forget to extend my thanks to my friends, Faten, Shayyma, Lara for their encouragement to me during my school years and to everyone who passed through my life and had an impact, even with a word, on what I am on.

Finally..

I would like to thank myself very much for the determination and dedication to work, for the ability to overcome all difficult circumstances.

TABLE OF CONTENTS

<u>Title</u>	<u>Page</u>
DEDICATION	I
ACKNOWLEDGMENT	II
TABLE OF CONTENTS	III
LIST OF FIGURES	V
LIST OF TABLES	VI
LIST OF APPENDICES	VII
LIST OF ABBREVIATIONS	VIII
ABSTRACT	IX
Chapter One: Introduction	1
1.1 Overview	1
1.2 Problem Statement	3
1.3 Aim of Study	4
1.4 Research Contributions	4
1.5 Thesis Structure	5
Chapter Two: Background & Literature Review	6
2.1 Background	6
2.2 Literature Review	7
2.3 Semantic Segmentation	10
Chapter Three: Jordan Forests Dataset	11
3.1 Dataset Collection	11
3.2 Ground Truth Dataset	17
3.3 Dataset Statistics	18
Chapter Four: Methodology	20
4.1 The Proposed Framework	20

<u>Title</u>	<u>Page</u>
4.2 Dataset Preparation	22
4.3 Image Preprocessing	22
4.4 Deep learning Architectures	23
4.4.1 U-Net	23
4.4.2 Initialization Models	25
4.5 Feature Extraction and Semantic Segmentation	25
4.6 Deforestation Monitoring	26
Chapter Five: Results and Discussions	29
5.1 Experiments Configuration	29
5.1.1 Environment Setup	29
5.1.2 Programming Language and Libraries	29
5.1.2.1 Python for Data Science	30
5.1.2.2 PyTorch	30
5.1.3 Hyper-parameters Optimization	30
5.1.4 Evaluation Metrics	31
5.2 The Results of Semantic Segmentation	32
5.3 Deforestation Monitoring Performance	34
Chapter Six: Conclusions and Future Work	38
6.1 Conclusions	38
6.2 Future Work	39
Appendix One: Specifications of Pretrained CNNs Architectures	40
References	43
Abstract in Arabic Language	48

LIST OF FIGURES

<u>Figure</u>	<u>Description</u>	<u>Page</u>
1	Remote sensing ways [1].	7
2	(a). Image at 1.17Km, (b). Image at 2.33Km.	12
3	Sample of an image with cloud.	12
4	Sample forests taken from Irbid, Al>Showbak and Zayy between 2010 and 2019.	13
5	Sample forests taken from Ajloun between 2010 and 2019.	14
6	Sample forests taken from Amman between 2010 and 2019.	15
7	Sample forests taken from Jerash between 2010 and 2019.	16
8	A sample ground-truth image drawn by AWS.	17
9	A set of original images with their corresponding masks.	18
10	A depiction of the proposed framework for Deforestation Monitoring.	21
11	Mask image over the original image.	22
12	U-Net architecture[2].	24
13	Semantic Segmentation.	26
14	The result of testing phase.	27
15	Smample of predicted images of Ain-Albash region.	28
16	The accuracy and the loss of Mobilenet_V2 model.	32
17	The accuracy and the loss of VGG16 model.	33
18	The accuracy and the loss of Resnet50 model.	33
19	The MIoU and the accuracy of base models.	34
20	Sample of changes in various regions around the years.	35
21	MobileNetV2 blocks architecture [3].	40
22	VGG16 architecture [4].	41
23	Resnet50 architecture [5].	42

LIST OF TABLES

<u>Table</u>	<u>Description</u>	<u>Page</u>
5	A Summary of recent works devoted for deforestation monitoring.	9
6	Forest distribution regions and the number of images taken from each region.	19
7	The characteristics of baseline CNN architectures.	25
8	Configuration of hardware and software	29
9	Hyper parameters used in model initialization.	31
10	The accuracy of base models using 30 epochs.	34
11	The MIoU of base models using 30 epochs.	34
12	The loss and gain rates over different years in 5 forest regions.	37

LIST OF APPENDICES

2cm

<u>Appendix</u>	<u>Description</u>	<u>Page</u>
One	Specifications of Pretrained CNNs Architectures	40

LIST OF ABBREVIATIONS

<u>Abbreviation</u>	<u>Description</u>
ML	Machine Learning
CNN	Convolutional Neural Network
DCNN	Deep Convolutional Neural Network
MIoU	Mean Intersection Over Union
std	standard deviation
AI	Artificial Intelligence
GFW	Global Forest Watch
EF	Early Fusion
SN	Siamese Network
CSVM	Convolutional Support Vector Machine
Te	Training epoch
Opt	Optimizer
Bs	Batch size
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
WCE	Wireless Capsule Endoscopy
GI	Gastrointestinal Diseases
SVM	Support Vector Machine
PFE-UNet	Pyramid Feature Extraction -UNetwork

ABSTRACT

MONITORING DEFORESTATION IN JORDAN USING SEMANTIC SEGMENTATION AND DEEP LEARNING

By

Lujain Hamed Mohammad Al-Smadi

Jordan is witnessing huge transformations in the nature of the environment and topographic features, where the desert represents a large part of Jordan lands with a very limited forest area. Over the last three decades, Jordan has lost about a third of natural forests at a rate of 1.6% annually. To help in monitoring deforestation in Jordan, we have collected a large dataset of Jordan forests in form of high-resolution satellite images. The main aim of this thesis work is to develop a deep learning model to automatically monitor deforestation based on semantic segmentation. It's worth mentioning that there is a lack of studies working on monitoring deforestation using semantic segmentation with deep neural networks. Therefore, we contribute in collecting a new dataset of Jordan forests over a range of ten years, as well as providing an accurate AI-enabled solution for monitoring deforestation in Jordan. The proposed model includes four essential stages: data preprocessing, deep architecture initialization, feature extraction, and semantic segmentation. We employed the efficient U-Net architecture, i.e., encoder/decoder model, to extract a set of discriminating features for semantic segmentation in order to predict the deforestation in the test images. The corresponding ground-truth images are also prepared with careful consideration to the forest masks drawn for each forest. Additionally, we initialized and examined our architecture by one of three pre-trained models: MobilenetV2, VGG-16 and Resnet50. These deep CNN-based models are fine-tuned on the domain of forest images through the procedure of transfer learning from the domain of general-purpose images. The performance of the semantic segmentation process has shown a high capability in detecting the forest and predict the surrounding areas. Moreover, the experimental results proved the model effectiveness in predicting the loss and gain rates in forests by reporting accuracy of 94.8%, and MIoU of 82% for Mobilenet_V2. The proposed model also showed a comparable accuracy using VGG16 and Resnet50 by achieving 94.7% and 94.3% respectively. The proposed model has succeeded in identifying the forest affected by any kind of deforestation over a range of years, and it is able to estimate the percentage of forest change, i.e. gain or loss, after applying a similarity check on the forests under consideration.

Chapter One: Introduction

1.1 Overview

Forests span a vast area of different terrain which include trees, herbs, plants, rivers and lakes. It hosts many animals algae, bacteria and other living organisms, and it is considered a treasure and an essential part of the global ecosystem. Forests are of great importance according to their beneficial effect on the climate. The presence of forests in an area makes it more moderate in temperature and more humid than the areas without forests, as well as absorbing large amounts of gases and various harmful air pollutants from the atmosphere.

Around 30% of Earth's surface is covered by forests. Food, fuel and medicine are provided for billions of people by these forests. Moreover, millions of people work in the forest sector. For instance, the Amazon rainforest covers an area of 5.6 million km², which is half of the surviving tropical forest area on the space [6][7]. It is home to a huge collection of animals and plants. Additionally, 20% of free-flowing freshwater comes from the Amazon river, it is the water source for most of South America's areas. Amazon rain forests help in local and global climate stability¹.

According to the National Institute for Space Research, during the 1990s, deforestation accelerated significantly in the Amazon in Brazil [8], and if the deforestation rate continues, Amazon rain forest will vanish by 2030 [9]. A recent study [10] highlighted the risks related to anthropogenic pressure on Ukraine forests. In 2019, illegal logging increased to 100, 000 cubic meters.

Over the years, tree removal increased significantly in various ways. Some of these operations were made by humans for construction or energy purposes, or without human intervention, e.g., natural disasters. Consequently, the impact of deforestation is obviously contributing the climate change and emissions of greenhouse gas and is one of the largest sources of CO₂ emissions [11].

Jordan has been considered one of the poor countries in terms of forest resources even that in the recent past it covered vast areas of it, up to Azraq, as evidenced by the remains

¹<https://www.worldwildlife.org/threats/deforestation-and-forest-degradation>

of oak trees in the Safawi region. The forests in Jordan suffered from threats that led to the reduction of the forest area, which today amounts to about 0.9% of the total area of Jordan Kingdom, and spreads from Wadi Al-Yarmouk in the north to Wadi Musa and the Sharah heights in the south. This percentage includes the forests that the Ministry of Agriculture², represented by the Forestry Directorate, reforested, as it helped reduce the decline of the forest area.

In Jordan, there are currently about 460,000 dunums of artificial forests, about 378,000 dunums of natural forests, and about 60,000 dunums of private-owned forests. There are plant patterns in Jordan, such as oak forests or evergreen oak, which are represented in many areas, for example, oaks are found in Shobak, Dana, Dabouq and Wadi Al-Safsa, and we find it in the north in Jerash, especially the high areas. The Ajloun Protected Area³ is considered a pure presence of oaks. Another pattern is the Aleppo pine forests located mainly in the Dibeen Forest Reserve area, in addition to the secondary communities in Ajloun, Suf, Mahis, Fuheis and Zi. Dibben Reserve establishing to preserve Aleppo pine forests, and the majority of reforestation projects use Aleppo pine because it is fast-growing. Among other plant types, the oak forest or the punitive oak is considered the national tree of Jordan, and it is one of the types of deciduous oaks and is represented in many regions of the Kingdom and its density reaches 60% and sometimes 70%.

Forest monitoring has been largely studied due to its essential impact on human life in many perspectives. Many strict policies and global laws have been put in place to monitor and protect forests with penalties. With the remarkable advances in Artificial Intelligence (AI) technologies, several platforms have been developed to monitor forests and take high-resolution satellite images such as Planet [12] and Global Forest Watch (GFW) [13]. In order to make forests data more reliable, AI has been used by linking it to crowdsourcing. GFW [13] is a large platform that captures millions of satellite images with the help of crowdsourcing, it helps conservation organizations by counting trees in ways that tracking illegal deforestation and enhanced forest management.

Recently, deep learning has attracted the researcher's attention in remote sensing because of its ability in extracting features from the image dataset automatically, mapping in complex environments [14] and high-level semantic segmentation that is recently used in various tasks. High potential in remote sensing has been shown through using deep learning. This led to producing state-of-the-art results in land-use and land-cover classification [15], object detection [16], semantic segmentation [17]. Recently, various works used deep learning for deforested areas detection in regions of the Brazilian Amazon and the result was satisfactory.

²<https://portal.jordan.gov.jo/>

³<http://www.wildjordan.com/ar/destinations/ajloun-forest-reserve>

Semantic segmentation is one of the challenging tasks in computer vision, and it plays an essential role in image analysis and understanding. It has various applications in artificial intelligence and computer vision such as robot navigation, medical imaging, and remote sensing. Recent works that utilize deep learning models have been improved significantly to deal with semantic segmentation using many efficient deep neural networks, thereby providing a faster and more precise segmentation. However, there is a lack of studies that examine the semantic segmentation approach in satellite imagery for deforestation monitoring. Therefore, we attempt in this work to employ deep learning with semantic segmentation to monitor deforestation in Jordan over a range of years automatically. This task requires collecting a sufficient amount of forest images that represent the Jordanian areas, which enable us to evaluate the proposed model.

1.2 Problem Statement

Forests are an important part of our nature, they are considered the lung of our life. Drought is one of the problems affecting the forests and vegetation, which is mainly related to climate changes. Fires also play a crucial role in increasing the deforestation spaces. Forests also suffer from human intervention, including overgrazing, whether in small afforestation projects or in large forests, and the loggers who use trees for general daily-life purposes. Forests also suffer from the premises located within the boundaries of forests in Jordan, especially in the intermittent areas, where people dispose and burn the wastes causing fires in the surrounding forests.

Deforested areas are usually investigated by forestry experts. But sometimes the field studies are hard in remote areas. Therefore, remote sensing is essential to make studies in these areas affected by deforestation. Satellites are able to remotely capture thousands of hectares on a single shoot with resolution arrived at 30 meters per pixel [18], and it is used significantly for capturing the forested areas. Therefore, many platforms have been provided to offer high-resolution satellite images, such as Planet [12] and GFW [13].

Jordan is considered one of the poor countries in terms of forests areas where the deterioration of green lands is increasing significantly. Therefore, there is a demand to monitor and study the deforestation that occurred over years in Jordan. The challenge of forests monitoring has become more reliable with the use of deep learning methods in which a precise analysis is performed for each pixel in the image. However, the accurate detection of forests in satellite images is not that simple task thus it is necessary to employ an effective segmentation approach to improve the model accuracy.

Moreover, any deep learning model developed with the aim of addressing the problem of deforestation monitoring is always restricted with the availability of a sufficient amount of images needed for training, validation and testing purposes. The quality of such image collections largely influences the performance of any intelligent methods dedicated to monitor deforestation. Consequently, we consider collecting a bench-marking dataset for Jordan forests to evaluate the proposed deep learning model for deforestation monitoring. As a result, we form our model framework by considering these problems and challenges, i.e., lack of local forest dataset, developing an accurate semantic segmentation approach, and learning a set of discriminating features through an efficient deep learning model.

1.3 Aim of Study

The main aim of this work is to monitor deforestation in Jordan using an efficient and accurate deep learning model in which the semantic segmentation of forest images is the core component. To achieve this goal, we aim to accomplish the following objectives:

1. Acquiring a collection of high-quality satellite images for all Jordan regions.
2. Creating the corresponding ground-truth collection by providing a set of forest masks for the original images.
3. Investigating the capability of semantic segmentation in identifying the forest areas and borders in any given image.
4. Monitor the deforestation in Jordan over a range of years to decide any changes in terms of forest gain or loss.
5. Learning the key features of forests through an efficient deep learning model to detect and analyze any forest changes which help in taking proper decisions.

1.4 Research Contributions

The main contribution of this thesis work can be summarized as follows:

1. New dataset of Jordan forests: We have collected a new dataset of forests images and landscapes in Jordan. This dataset represents many regions over many years to monitor any changes in the forest coverage areas. Also, we have prepared the ground-truth images of the original dataset which are annotated and associated with synthetic forest masks.
2. Semantic segmentation: We semantically segment the forest areas and any region of interest in satellite images. This process helps in identifying deforestation that has been

monitored over a particular period. Employing and effective semantic segmentation will also help in identifying any gain changes that occurred to the forests.

3. Deep learning model: We develop a deep learning model that includes a set of effective neural networks trained end-to-end in order to learn the discriminating features of forest images. These architectures consist of image encoders and decoders, as well as pre-trained ones for fine-tuning and optimization purposes.

1.5 Thesis Structure

Chapter 1 provides a general introduction to deforestation monitoring, including the research problem, aim of the study, and thesis contributions.

Chapter 2 introduces a background on deforestation and semantic segmentation, and it reviews the most related works devoted for deforestation monitoring.

Chapter 3 presents a new dataset of Jordan forests and the procedure of image collection and preparation.

Chapter 4 discusses our methodology and illustrates the proposed deep Convolutional Neural Network (CNN) based model for deforestation monitoring.

Chapter 5 explains the experiments setups and the performance results of our proposed model.

Chapter 6 concludes this thesis and summarizes the main findings and how they relate to our research problem and objectives. It also outlines the possible directions that could expand this research work in the future.

Chapter Two: Background & Literature Review

2.1 Background

Deforestation indicates a decrease in forest areas which missing for other uses such as urbanization, agricultural croplands, or mining activities. Human activities are one of the main factors of deforestation and the expansion of agriculture caused nearly 80% of global deforestation [19]. Deforestation has been negatively affecting climate, biodiversity, and natural ecosystems. So it is necessary to monitor the forests and protect them from accelerating human activities so that the problem does not escalate and become more complex.

Remote sensing [20][21][22][23], as shown in Figure 1, is a monitoring and detecting process of an area at a distance from an aircraft or satellite. Special cameras on satellites take images of the Earth's surface which allows us to see more regions especially inaccessible and remote landscapes such as forests. Satellite images help to keep tracking the land changes such as forest fires [24], deforestation and weather monitoring which helps to prevent various bad reflected results.

Satellites like Sentinel and Landsat are able to capture with one shoot (scene) several thousands of hectares with a resolution of 10 to 30 meters per pixel image. This technology also outperforms the capability of standard drones that might only capture up to 1.000–3.000 hectares per day. Remotely sensed imagery provides a synoptic view of the environment in its view range. Therefore, satellites have been utilised in many fields, especially in forest monitoring, e.g., forest loss or gain.

Recently, deep learning has come to solve various problems in various areas. In the medical domain, Ouahabi et al. [25], proposed a method in order to improve the efficiency of segmentation while keeping high accuracy. Gulshan et al. [26] employed a Deep Convolutional Neural Network (DCNN) on an eye picture archive communication system dataset, to classify and detect moderate and worse referable cases. Jia et al. [27], also applied DCNN on wireless capsule endoscopy(WCE) images for bleeding detection in Gastrointestinal Diseases(GI). In computer vision, deep learning has been implied to solve various problems.

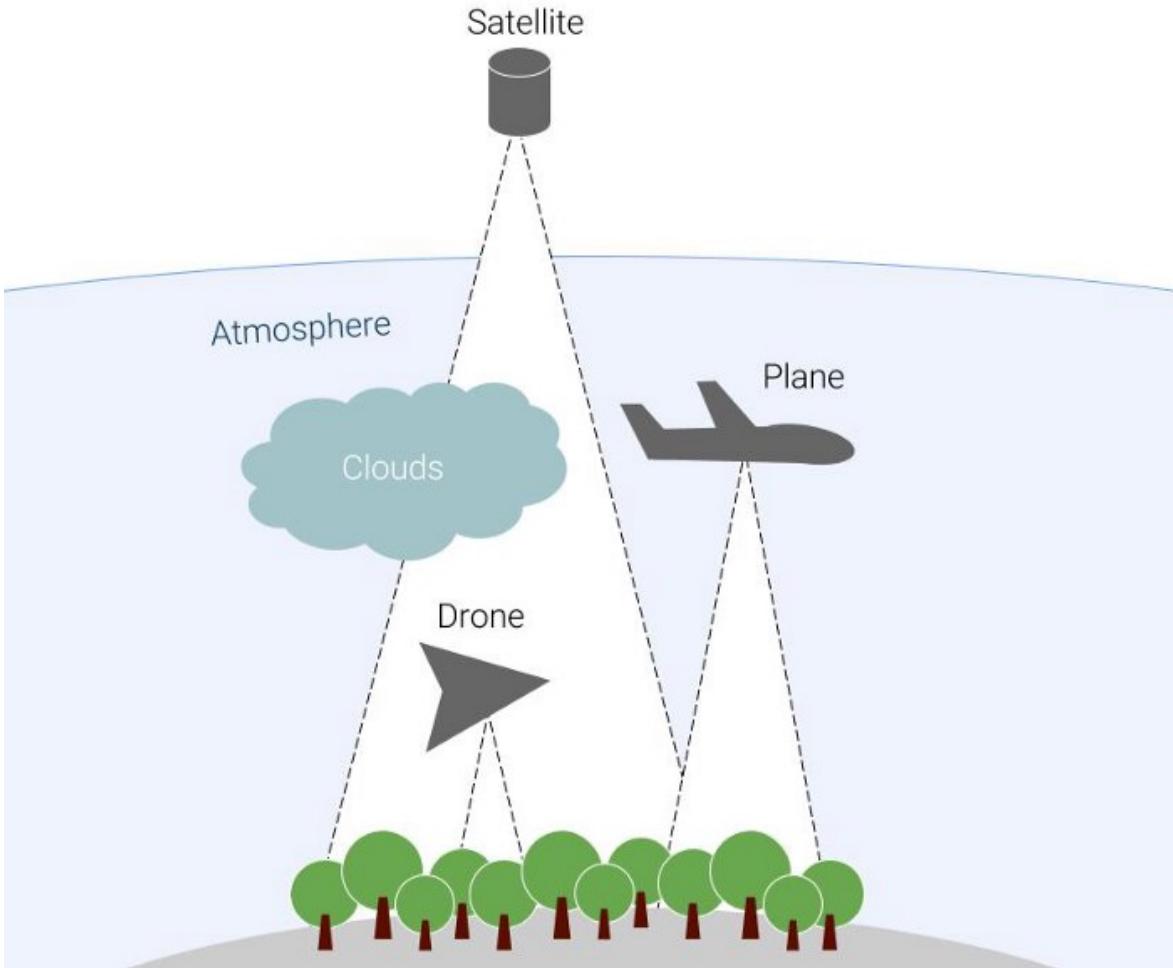


Figure 1: Remote sensing ways [1].

Face recognition is one of the hottest computer vision applications with high commercial interest. Lawrence et al. [28] employed a CNN for face recognition.

The remaining part of this chapter is organized as follows: Section 2.2 presents an overview of deep learning methods related to deforestation monitoring, Section 2.3 provides a background of semantic segmentation.

2.2 Literature Review

Recently, researchers' interest in vegetation has increased significantly. Mhatre et al. [29], proposed a CNN-based approach to monitoring the vegetarian distribution using satellite images. The CNN model was trained by using supervised learning to get an accurate determination of forest cover. Rakshit et al. [30] used a classification VGG16 model to track the changing land pattern in the Amazon rainforests. Bragagnolo et al. [31] evaluated U-Net

performance for forest mapping. This study used Sentinel-2 satellite imagery for the Amazon forest. The result showed that U-Net is able to identify and draw forest areas polygons. Chantharaj et al. [32] presented a modified SegNet framework in addition to a method for adding more feature bands into satellite images.

The studies increased in deforestation domain mainly on Amazon rainforest. Brovelli et al. [33] concentrated on mapping and monitoring the changes in the forest from 2019 and 2020, in addition to simulating the development of future rainforests in Brazil. Depending on a supervised classification model available on the Google Earth engine. Machine learning played as the core of this research, the dynamics of land cover were determined at 5-year intervals. The results allowed estimating the deforestation rate over a long period and analysing the forest dynamics. Wyniawskyj et al. [34], proposed a method for forest monitoring in real-time to help in preventing any illegal activities such as logging activities and forest fire scars. Convolution2D, Max-Pooling2D and Dropout have trained on a satellite image in the period between 2017 and 2018.

Similarly, Ortega et al. [35] worked on deforestation monitoring using CNN-based models and the PRODES database. Siamese CNN (S-CNN), Early Fusion (EF), and the Support Vector Machine (SVM) as a baseline were tested. S-CNN was superior in terms of F1-score and overall accuracy to its counterparts, and it identified deforested areas. Whereas EF was better than SVM that gave a weak performance for deforestation detection. Another study using FCN-based was introduced in [36], where SegNet and U-Net algorithms showed low performance for land use types with a small number of pixels, and within non-forest lands, a misclassification occurred when dividing the U-Net results.

Maretto et al. [37] proposed Spatio-Temporal deep learning in order to map deforestation in the Amazon rainforest using U-Net architecture to help in combining both temporal and spatial contexts. This method achieved 95% accuracy using SAR images taken for the wildfires and CNN has been developed to identify burnt zones [38].

Isaienkov et al. [10] presented a baseline U-Net model for detection of Ukraine forest deforestation. The training is done on a custom dataset created on Sentinel-2 imagery and they used ArcGIS for labelling. Andrade et al. [39] evaluated and implemented a method of deforestation detection. This method depended on a Fully Convolutional Deep Learning model. The study has concentrated on Landsat OLI-8 images in 2016 and 2017 for the Amazon region and outperformed the S-CNN and Early Fusion methods in terms of F1-score.

Table 5 represents a comprehensive summary of all research related to forest monitoring with regard to deforestation, the regions that are covered by this research, and methods used in these works.

Table 5: A Summary of recent works devoted for deforestation monitoring.

Author	Dataset	Method/Model	Spatial Resolution	Accuracy	Region
Brovelli et al. [33], 2020.	Landsat, Sentinel-2 data	ML supervised classification	10-30m	90-92.5%	Brazil
Andrade et al. [39], 2020.	Landsat OLI-8	DeepLabv3+	30m	97%	Brazil
Bragagnolo et al. [31], 2021.	Sentinel-2 satellite imagery	U-Net	10m	94.70%	Brazil
Rakshit et al. [30], 2018.	Data collected from Planet	CNN	3.7m	96%	Brazil
Ortega et al. [35], 2019.	PRODES	EF, SN, CSVM	30m	90%	Brazil
Maretto et al. [37], 2021.	OLI-8	Spatio-temporal U-Net	30m	95%	Brazil
Wyniawskyj et al. [34], 2019.	Sentinel-2 optical imagery, Digital Globe and Planet platforms	Shallow Neural Network	10m-1k	N/A	Guatemalan
Ban et al. [38], 2020.	Sentinel-1 SAR data	CNN	N/A	83%	Canada,USA
Isaienkov et al. [10], 2021.	Sentinel-2 image	U-Net	N/A	N/A	Ukraine
Chantharaj et al. [32], 2018.	Landsat-8	SegNet	15m	80%	Thailand
Lee et al. [36], 2020.	Komsat-3 images	SegNet and U-Net	0.7m	74.8%	Korea
Mhatre et al. [29], 2019.	Kaggle's Dataset	CNN	3m	90%	India

2.3 Semantic Segmentation

The advances of deep learning models, mainly DCNNs, led to a notable improvement in the field of semantic segmentation. The aim of semantic segmentation is to assign a label for each pixel in the image. Semantic segmentation have been applied in many applications such as medical imaging [40], image retrieval [41][42] , and object detection [43].

However, there are few research works that have employed semantic segmentation in monitoring deforestation. Zulfiqar et al. [44] proposed a DCNN-based approach that uses semantic segmentation to process satellite images in order to monitor the forest change patterns. This work has been performed using Landsat-8 multi-spectral imagery from 2014 to 2020 and generated pixel-wise forest and non-forest change maps. Costa et al. [45] investigated the ability of deep semantic segmentation in recognising eucalyptus afforestation areas in Brazil. Six architectures were compared with 4 encoders by which an IoU of 76.57 is achieved.

Wang et al. [46] conducted a study to investigate the optimal segmentation algorithm in addition to spatial resolution for eastern redcedar. The best accuracy achieved was 0.918 obtained by the ResNet model. Zhang et al. [47] also proposed a Pyramid Feature Extraction-UNet network (PFE-UNet) for forest image segmentation. This study used the DeepGlobe dataset. The accuracy obtained was 94.2%. Garg et al. [48] examined three traditional machine learning algorithms against a deep learning model for semantic segmentation, and they showed that the pre-trained deep learning model DeepLabv3+ exceeds the conventional algorithms for land cover and land use.

Our work is distinguished from others in that we have built a new data set for the forests of Jordan. We extracted deep and compact features using semantic segmentation and deep learning, where we used a U-Net architecture with three base models. Furthermore, we were able to determine the loss and gain through a procedure of similarity check between the semantic segmentation of images over many years.

Chapter Three: Jordan Forests Dataset

This chapter presents the dataset that has been collected for Jordan forests. Section 3.1 presents the dataset properties and the regions that have been studied and some samples from these regions; Section 3.2 presents the ground truth dataset and its properties in addition to some samples that have been drawn; Section 3.3 presents dataset statistics.

3.1 Dataset Collection

We have collected a set of satellite images from the forests in Jordan, which spans all lands from north to south using Google Earth Pro¹. The images represent various regions mainly in the governorates of Ajloun, Irbid, Jerash, Amman and Shoubak. The number of collected images is 4,576 images. We made a comprehensive scanning of forest distribution areas over a ten-year period from 2010 to 2020. The images were taken over two heights, the first was at 2.33 km to include a large area of the forest, and the second was at 1.17 km, which includes a narrower area of the forest so that the particular forest can be divided into several parts, Figure 2 shows some sample images. As a result, hundreds of images of a single forest have been obtained over a period of 10 years as shown in Figure 4, 5, 6 and 7. The program provides an option to prevent the appearance of clouds so that the collected images are cloud-free with a high resolution of 4K, which is a very high resolution. Figure 3 shows an example of an image with clouds.

Figure 4 shows a sample of forests have been taken from Irbid governorate for Al-Nuayymah, Al-Korah and Almazar, in addition to Al-Shobak governorate and Zayy region. Whereas Figure 5 for Ajloun governorate regions; Ajloun Forest, Alwahadneh, Ain-Jannah, Anjarah, Dair-Smadieh, Ishtafena and Jdet. Figure 6 contains a sample taken from Amman governorate; Ain Albasha, National Park, Albahrain forests, Iraq ALameer, Mahis, Um Aljawz and Yadudah. Figure 7 for Jerash regions; Borma, Dibeen, Rajeb, Sawf, Thaghret Asfour and Zobia. All of these samples are between 2010 and 2019.

¹<https://www.google.com/earth/vendors/>



a.

b.

Figure 2: (a). Image at 1.17Km, (b). Image at 2.33Km.



Figure 3: Sample of an image with cloud.

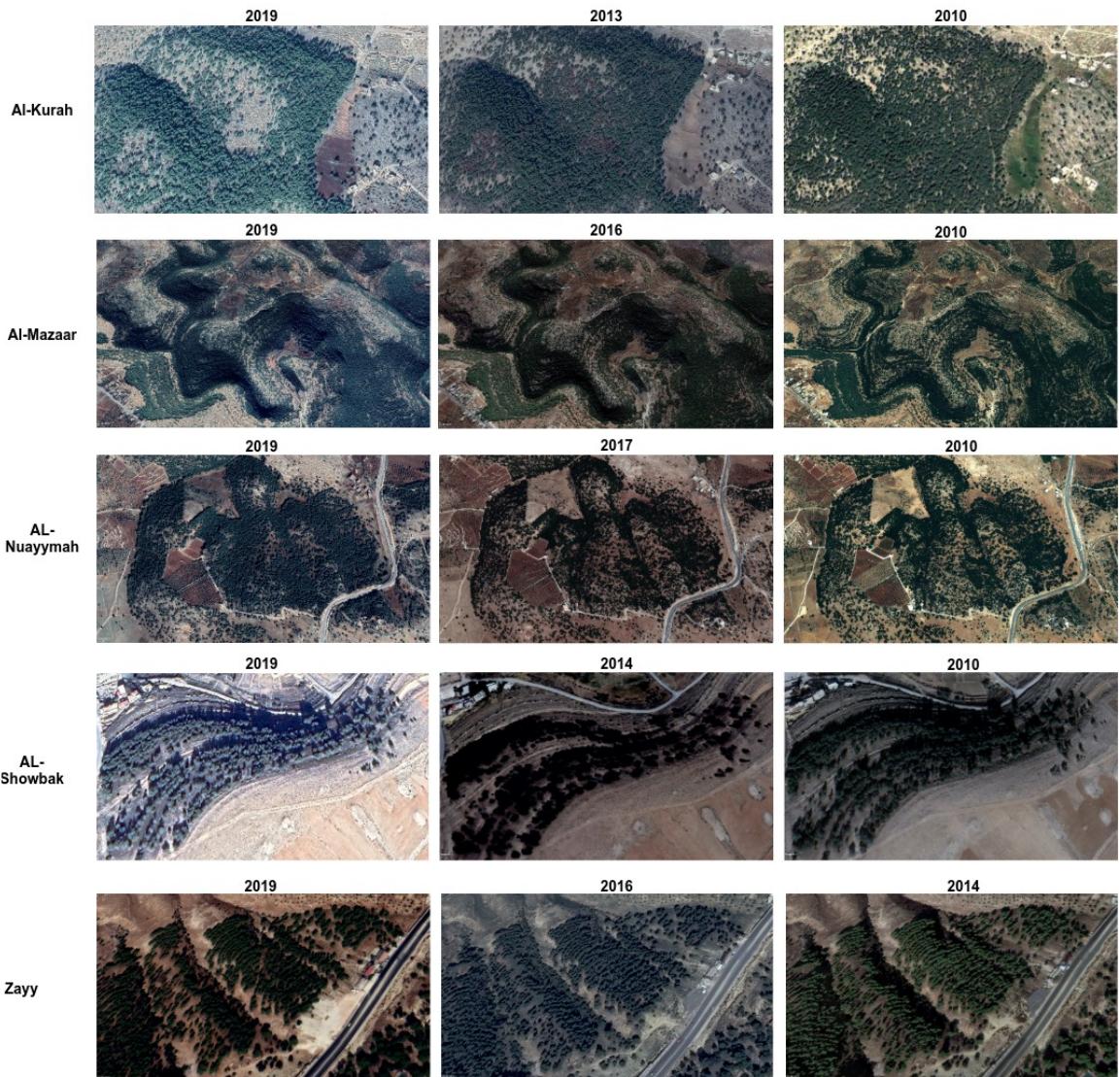


Figure 4: Sample forests taken from Irbid, Al-Showbak and Zayy between 2010 and 2019.

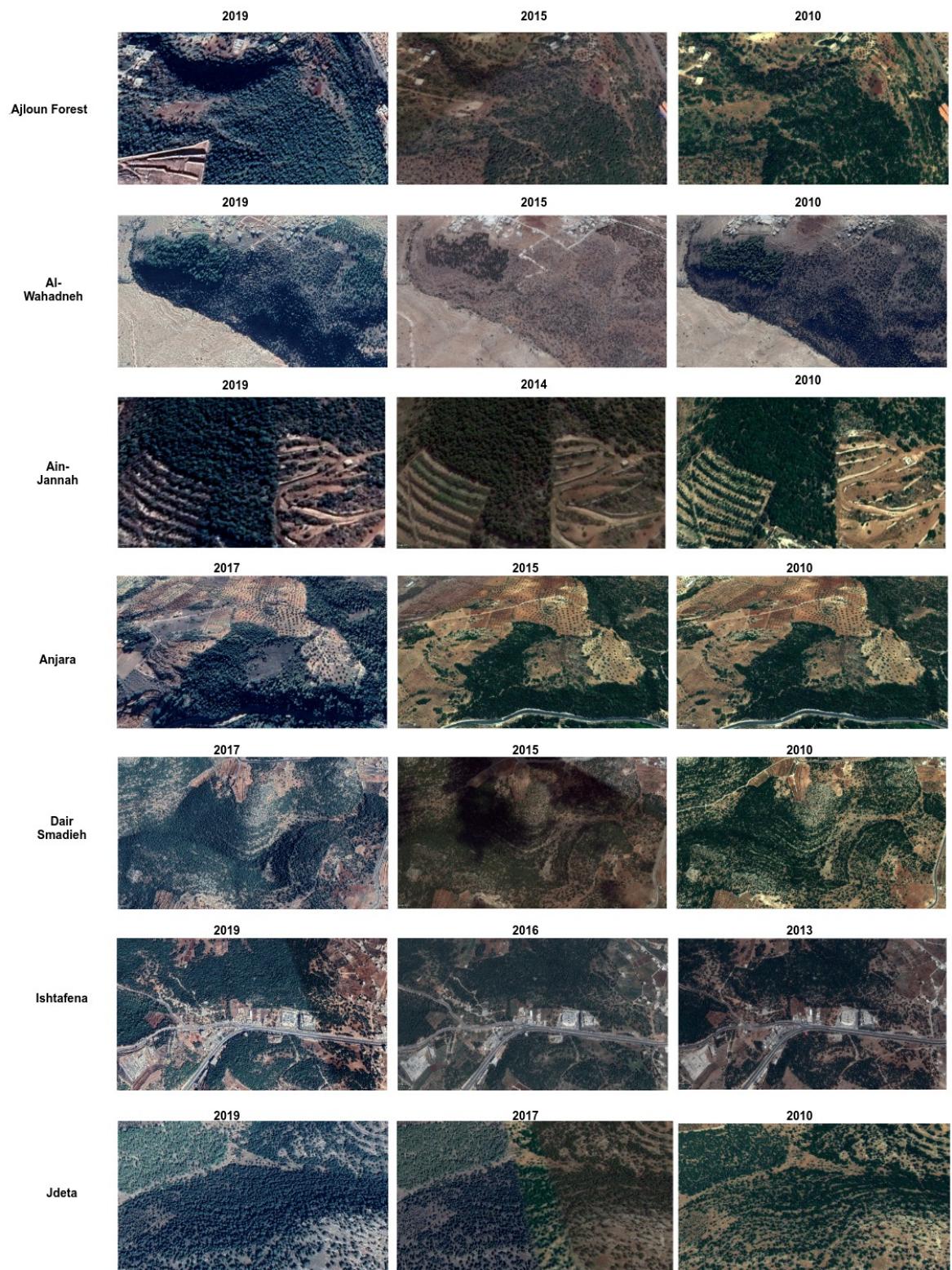


Figure 5: Sample forests taken from Ajloun between 2010 and 2019.

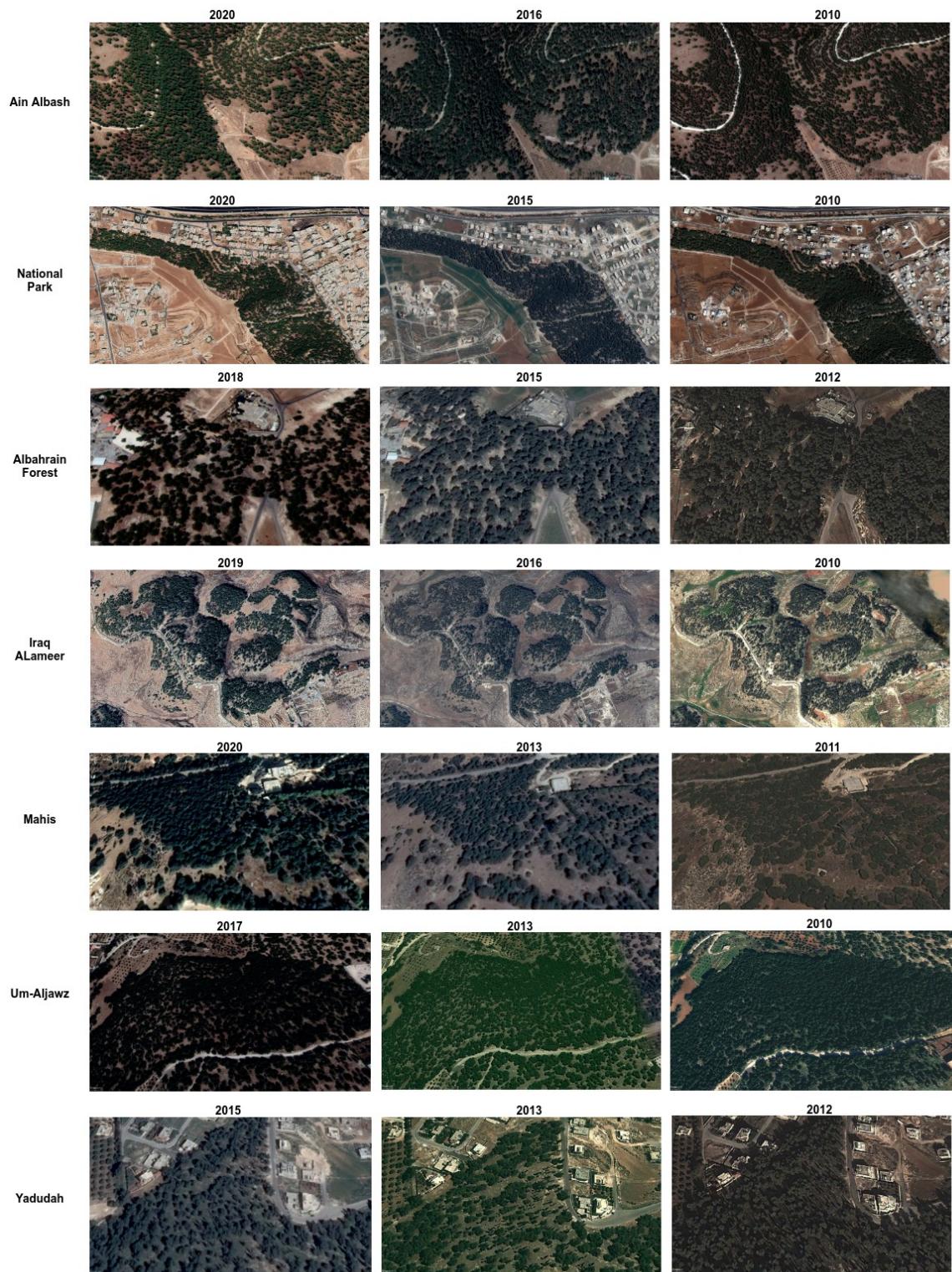


Figure 6: Sample forests taken from Amman between 2010 and 2019.

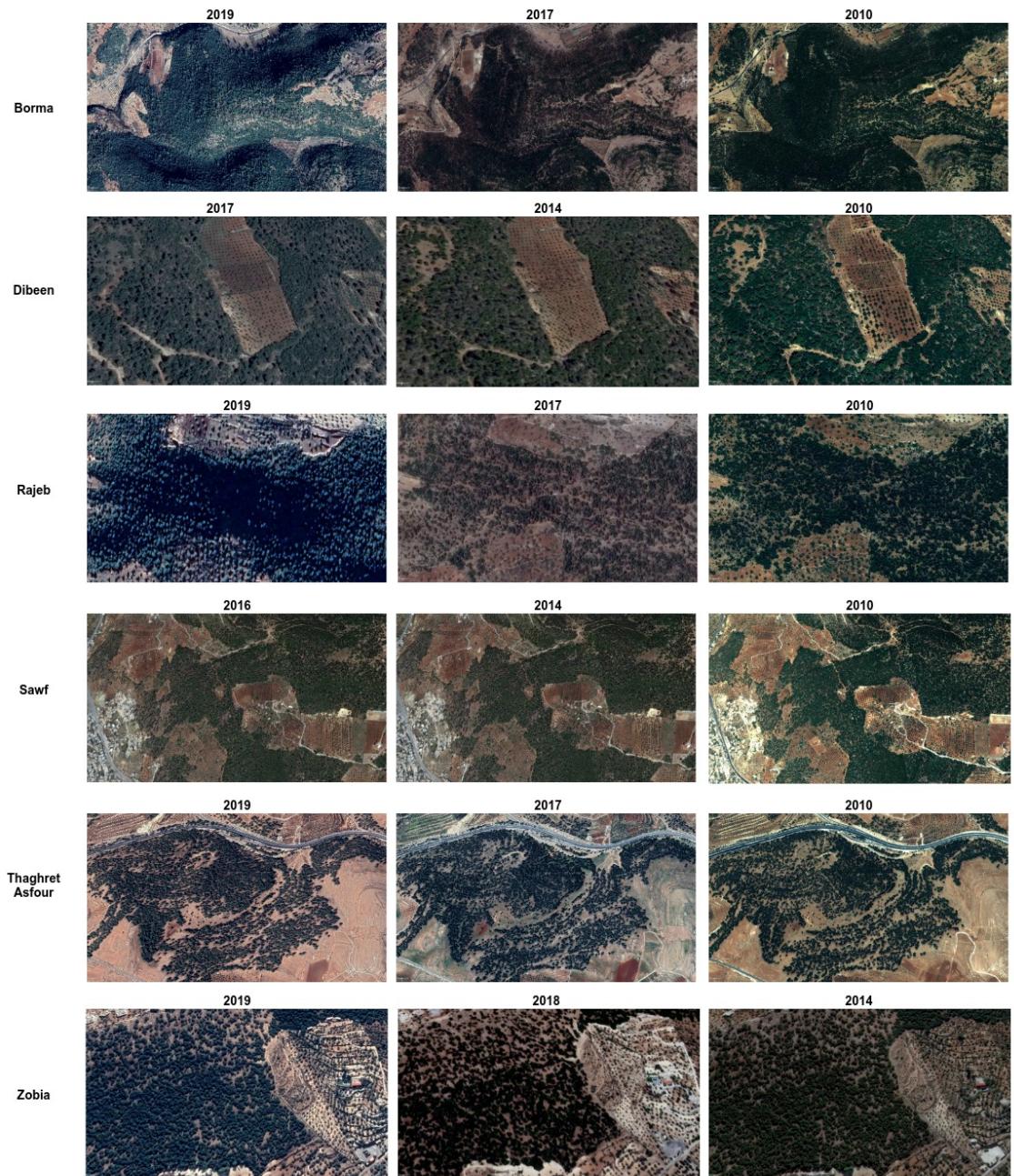


Figure 7: Sample forests taken from Jerash between 2010 and 2019.

3.2 Ground Truth Dataset

The ground truth of these images are prepared over Amazon Web Service(AWS). Plotting images was very difficult and required very high accuracy. The original images are stored with the extension (jpg) but the mask images are (png).

The ground truth images have two colors, we assigned the green color to represent the forests, and the other objects defined as a background in white color. Therefore, there are two classes in each image; forest and non-forest. The image was drawn so that the forest boundaries are surrounded by clicking on successive clicks starting with a point and ending with it as shown in Figure 8.

All the mask points are merged together forming the boundary of the forest. Then the forest is highlighted with a green mask and the remaining objects take white color to denote the background area as shown in Figure 9. We focused on the boundaries that are dense with trees, i.e., areas with scattered trees are excluded from the forest mask.



Figure 8: A sample ground-truth image drawn by AWS.



Figure 9: A set of original images with their corresponding masks.

3.3 Dataset Statistics

We have collected the satellite images of various regions in most of the Jordan governorates. Table 6 summarizes the number of images taken in each region of Jordan. It can be noted that most of the forests are distributed in Ajloun with 9 forest areas, 8 areas in Amman, and 6 areas in Jerash.

Table 6: Forest distribution regions and the number of images taken from each region.

Governorate	Region	Number of images
Ajloun	Ajloun Forests	309
	Al Wahadneh	11
	Anjarah	26
	Ayyn Janna	315
	Barqash	176
	Dair Al Smadieh	118
	Ishtafena	494
	Jdeta	32
		Total: 1481
Amman	Ain Albasha	505
	Al Bahrain forests	150
	Iraq Al Ameer	136
	Mahis	234
	Umm Al Ajouz	88
	Yadoudah	110
	Amman National Park	200
	Zayy	60
		Total: 1483
Jerash	Borma	404
	Dibeen	210
	Rajeb	19
	Sawf	179
	Thaghret Asfour	275
	Zobia	78
		Total: 1165
Irbid	Al Kourah	8
	Al Mazaar	138
	Al Nuayymah	284
		Total: 430
Al Shobak	Al Shobak	17
		Total: 17

Chapter Four: Methodology

This chapter presents the methodology and techniques used to develop our proposed model for deforestation monitoring. Section 4.1 covers a description of the proposed framework briefly; Section 4.2 present a description of the dataset preparation utilized for the deforestation monitoring model; Section 4.3 provides the pre-processing operations applied on the input images; Section 4.4 discusses the model initialization using three pre-trained CNN models; Section 4.5 presents the process to shape the final image descriptor features extraction and semantic segmentation stage; Section 4.6 presents the deforestation monitoring and it explains how the semantic segmentation is applied.

4.1 The Proposed Framework

In this work, we propose a deep learning CNN-based model for deforestation monitoring using image semantic segmentation, as shown in Figure 10. The complete framework consists of four main stages :

- 1) Dataset pre-processing to improve the image data that suppresses distortions and enhances the important features such as noise reduction, contrast, image scaling and colour space conversion.
- 2) Model initialization using pre-trained CNNs to conduct an efficient transfer learning from a general purpose images to a specific domain images, i.e., forest images. This will be done through training. and
- 3) Features extraction and semantic segmentation to segment the forest coverage area of any region of interest. This process helps in identifying deforestation that has been monitored over several years. The framework details are explained in the next subsections.
- 4) Forest monitoring over a range of years to decide any changes in terms of forest gain or loss.

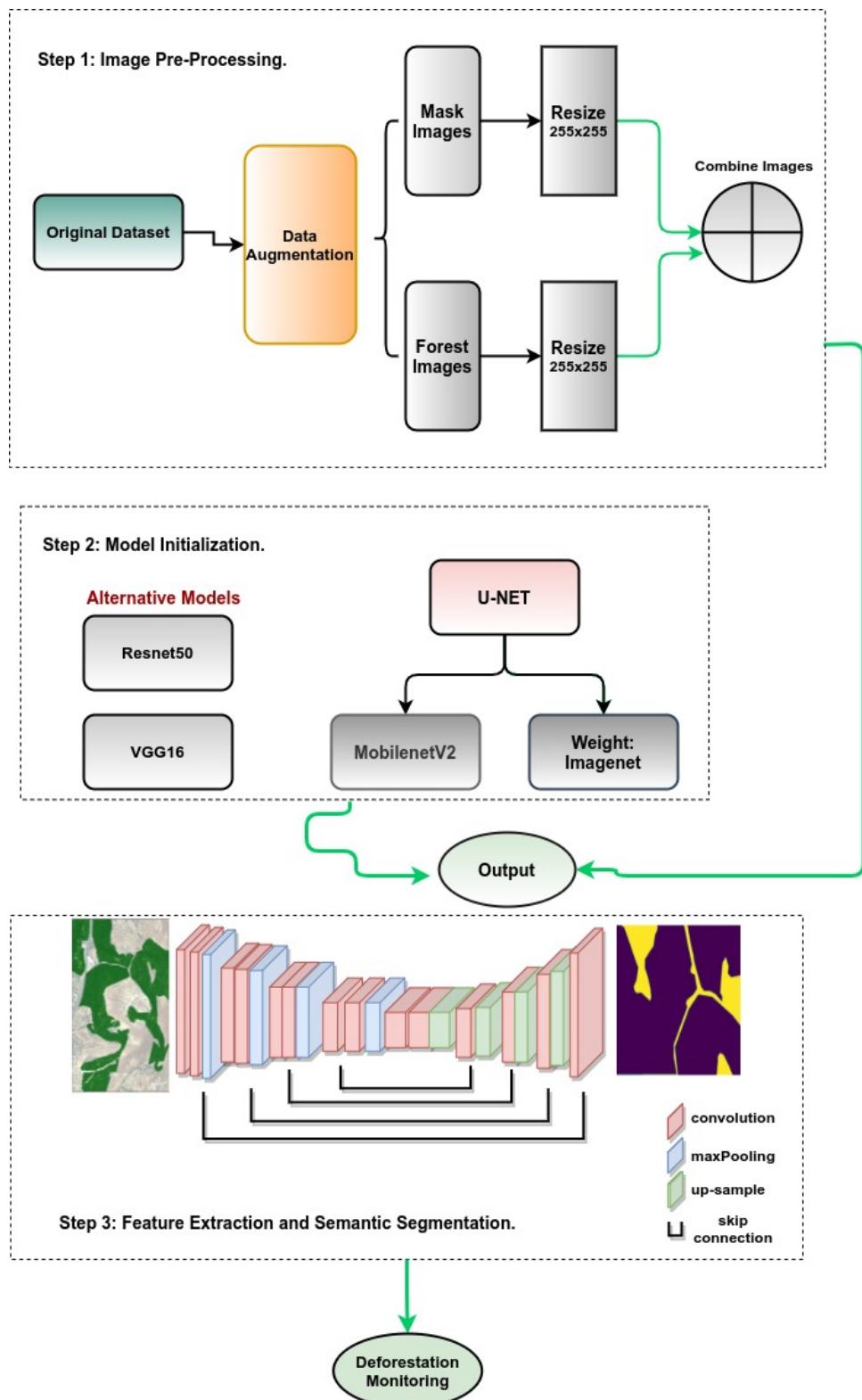


Figure 10: A depiction of the proposed framework for Deforestation Monitoring.

4.2 Dataset Preparation

We split the dataset into 3 parts: 70% for training, 20% for testing and 10% for validation. Therefore, 3294 images for the train set, 366 images for the validation set, and 912 images for the test set. The corresponding ground truth images are also have the same split procedure. Before feeding the data into the network, the mask image is applied over the original one to be as one image as shown in Figure 11.



Figure 11: Mask image over the original image.

4.3 Image Preprocessing

Image pre-processing is an important step to improve image quality that can help in better analysis. Enhancing features and suppressing undesired malformations can be done through this step. Image processing in our case has some steps:

1. Read Images : Reading images stored in the dataset path then transforming them into a numerical array.
2. Resizing images: The image resizing step establishes a base size for all inputs fed to the network. We resized all images to $(256 \times 256 \times 3)$.

Data Augmentation [49][50][51] is also important to learn the network desired robustness and invariance properties when there are a few samples for training available. Data augmentation is done before feeding data to the model. Data augmentation includes:

1. **Flip:** The image can be flipped vertically or horizontally. A vertical flip is valent to rotating the image by 180 degrees.

2. **Rotation:** After rotation, the image dimensions may be changed. For a square image, the image size will be preserved, but the rectangle image the size would be preserved if the rotation degree equals 180.
3. **Scale:** Images can be inward or outward. The outward scaling makes the size of the image larger than the original, while the inward scaling reduces the image size.
4. **Crop:** It is applied to a random sample section of an original image, then the chosen section will be changed to the original size.
5. **Translation:** It involves moving the images over X and Y directions or one of them.

In our work, we used vertical and horizontal flipping.

4.4 Deep learning Architectures

Deep learning and CNN have been highly ubiquitous in the domain of computer vision. CNNs are popular in various tasks of computer vision, such as image classification, image generation, object detection. This section introduces the deep learning architectures used in our work to initialize the baseline model and train the complete model.

4.4.1 U-Net

U-Net firstly used for biomedical image segmentation by Ronneberger et al. [52], as shown in Figure 12. The first half of the architecture is the encoder. The encoder is usually a pre-trained classification network such as VGG and ResNet. The encoder consists of the repeated application of two 3×3 convolutions followed by ReLU and 2×2 max-pool down-sampling with stride 2 in order to get feature representations at various levels through image encoding. For every downsampling the number of feature channels is doubled.

The second part of the U-net architecture is the decoder. The decoder consists of upsampling and concatenation followed by regular convolution operations. The goal is to offer discriminative features learnt by the encoder upon the pixel space (higher resolution) to acquire a dense classification. Skip connections allow gradients to flow better and grant information from multiple image scales. The information extracted from the upper layers improves classification accuracy, whereas those coming from deep layers supports the segmentation engine. A pixel-wise soft-max over the last feature map combined with the loss function (cross entropy) are used in order to compute energy function.

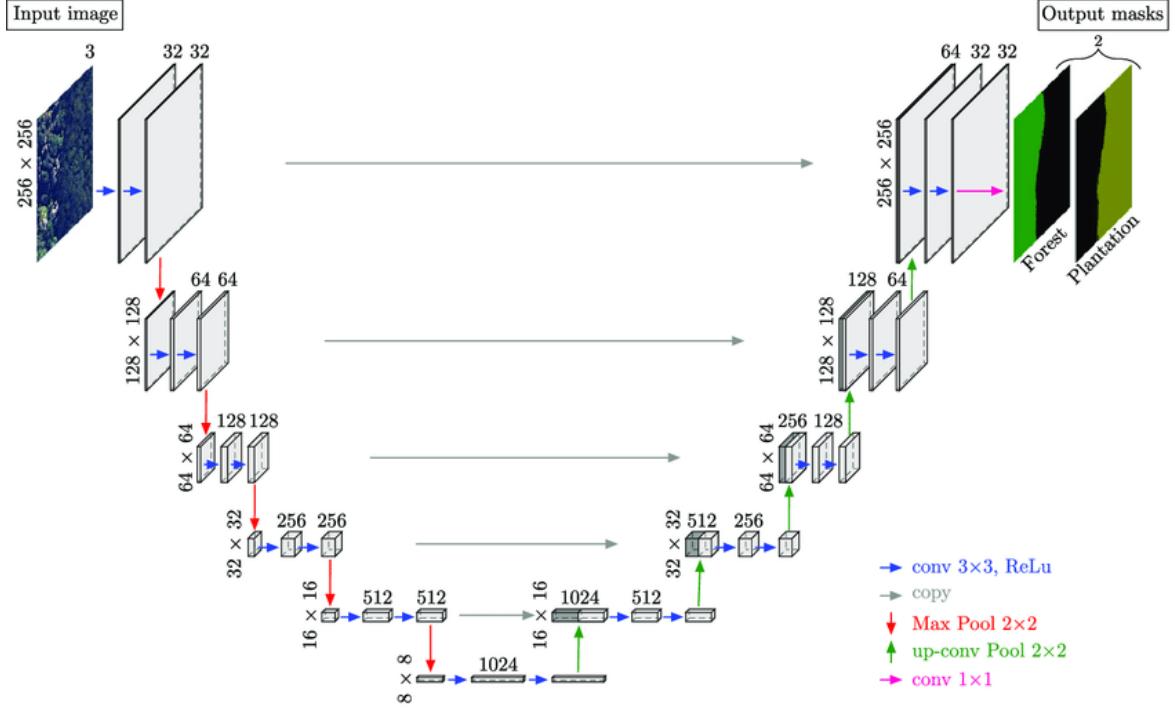


Figure 12: U-Net architecture[2].

Soft-max is defined as:

$$p_k(x) = \exp(\alpha_k(x)) / \left(\sum_{k'=1}^K \exp(\alpha_{k'}(x)) \right) \quad (4.1)$$

where α_k is an activation in feature channel $k' \in K$ at the pixel position $x \in \Omega$.

We used an input image of size $256 \times 256 \times 3$. In the encoder, the size of image is gradually reduced while the depth is gradually increased. After passing the image through two convolution layers and applying the max-pooling to reduce the image size, the number of layers is increased by doubling the number of filters for convolutional layers on each convolutional block. It starts from 32 filters on each convolution layer and 256×256 image size, then it increases the filters to 64 and decreases the image size to 128×128 for the second convolutional block. In the third convolutional block, the number of filters is 128 for each layer with an image size of 64×64 , then it is reduced to 32×32 in the fourth layer with 256 filters. The process continues until the image size of 8×8 and 1024 filters for each convolution layer are reached. These processes act as a feature extraction process. So, the aim here is to reduce the resolution and increase the depth, i.e., the number of layers.

Also, in the up-sampling path, we have a set of convolutional blocks with 2 convolutional layers in each. The number of filters for each consecutive convolutional block equals the half of the filters from a previous convolutional block. Starting from 1024 filters to 32 for each

convolutional layer. The aim is to increase the resolution and reduce the depth (number of layers). The skip connections are concatenating the feature maps from corresponding down-sampling layers for more precise localization of features detected in the blocks.

4.4.2 Initialization Models

Three well-performing CNN architectures are used in this work to initialize the baseline model with their optimal weights during the procedure of model fine-tuning. As summarized in Table 6, these architectures are VGG-16, ResNet50, and MobileNet.

- **VGG-16 [53]**

It includes 13 convolutional layers and fully connected layers. The activation function used in VGG16 is ReLU activation. VGG-16 contains 138 million parameters. The filters used here are small (2×2 and 3×3). It is an efficient architecture that shows high performance in many applications including object detection. VGG16 optimal weights are loaded after importing the Numpy module for array processing.

- **ResNET50 [54]**

It is based on skip connections or residuals in addition to using batch normalization. ResNet50 is a type of Resnet composed of 26 million parameters. It is used widely in object detection models and frameworks.

- **Mobilenet_V2 [55]**

It is based on a streamlined architecture to develop lightweight deep neural networks. It consists of 32 filters in addition to 19 residual bottleneck layers following them. The kernel size used here is (3×3) which is a standard for the modern network, and batch normalization and dropout are utilized [56].

Table 7: The characteristics of baseline CNN architectures.

Model	Total Layers	Parameters	Total Filters
VGG16	16	138 million	512
ResNet50	50	25.6 million	-
MobilenetV2	20	3.4 million	32

4.5 Feature Extraction and Semantic Segmentation

A set of deep learning models are provided and trained on the ImageNet dataset [57], and the CNNs used for extracting the key features are VGG16, ResNet and Mobilenet_V2. These networks are mainly employed in our segmentation and classification tasks.

For feature extraction, we used the output before the classification layer of VGG16, Resnet and Mobilenet_V2 models. For example, in VGG, the feature extraction part is produced by the last max-pooling layer, whereas the rest of the network is responsible for the classification part. The input image is loaded with the expected size 256×256 . Then, it converts object of the image by Python Imaging Library¹(PIL) to a Numpy array (pixel data), which is scaled to formulate the initial image features.

Semantic segmentation classifies each pixel in a given image from the predefined classes. The goal of semantic segmentation is to label the image pixels with the corresponding class of what is being represented, as shown in Figure 13. This task is usually considered as a dense prediction because the prediction occurs for every pixel in the image. Specifically, the goal is to take an RGB colour image with a size of $(\text{height} \times \text{width} \times 3)$ and generate a segmentation map where every pixel contains a class label performed as an integer $(\text{height} \times \text{width})$ [58].

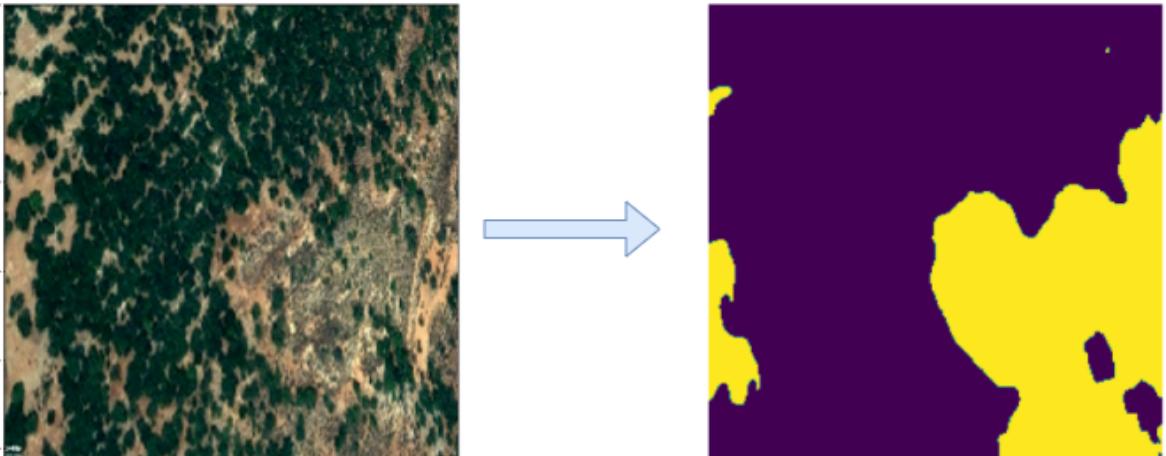


Figure 13: Semantic Segmentation.

It is worth to mention that the instances of the same class are not treated as separate parts. In particular, if we have two forests (two objects of the same category) in the same input image, it does not distinguish them as separate objects by the segmentation map [59].

4.6 Deforestation Monitoring

We evaluate the procedure of deforestation monitoring by feeding the proposed model with a set of test images to ensure the model effectiveness. Figure 14 shows the result after test stage, the predicted image is very similar to the ground truth.

In order to get the percentage of variation in forests over the years, we collect and save the predicted images into a separate directory. Then, we select the images that belong to the

¹<https://www.geeksforgeeks.org/python-pil-image-getdata/>

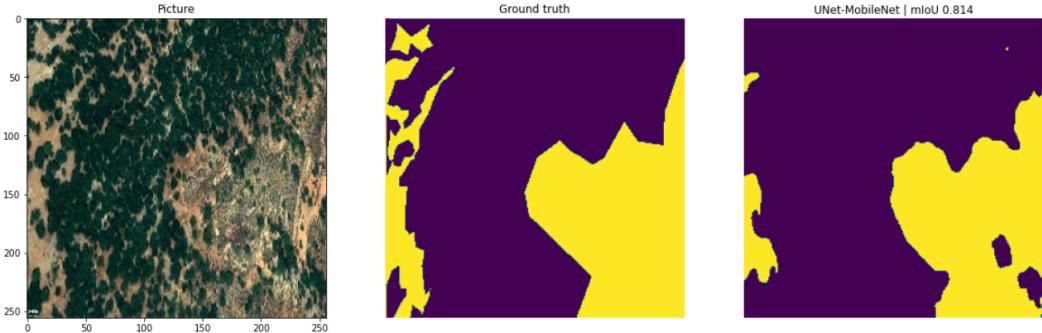


Figure 14: The result of testing phase.

same area but on different years, and we calculate the similarity rate in these images for the same area over different years. The final decision is either the forest has a loss or gain.

The collected images are the predicted images of the test set, which are 916 random images of different regions. We were able to get images of 5 areas over a range of about 10 years. The selection of images was performed manually. For instance, Figure 15 shows a sample of predicted images of different patches in the Ain-Albasha region over different years. The figure represents the changes occurred in the forests. These samples represent the resulting predictions of semantic segmentation performed on the test dataset.

We can distinguish and observe the change in the forests and find out which images have retreated or gained. In the next section, we made statistics on the areas where the forest has a gain or loss change.

We used Pillow’s ImageChops² to calculate the similarity/dissimilarity rates between the segmented images over different years. The imagecompare method from this library converts the input image to a grey-scale version, then it sum up all different pixels by summing up their histogram values. Finally, it calculates the percentage of differences based on a black-white image of the same size.

²<https://pillow.readthedocs.io/en/stable/reference/ImageChops.html>

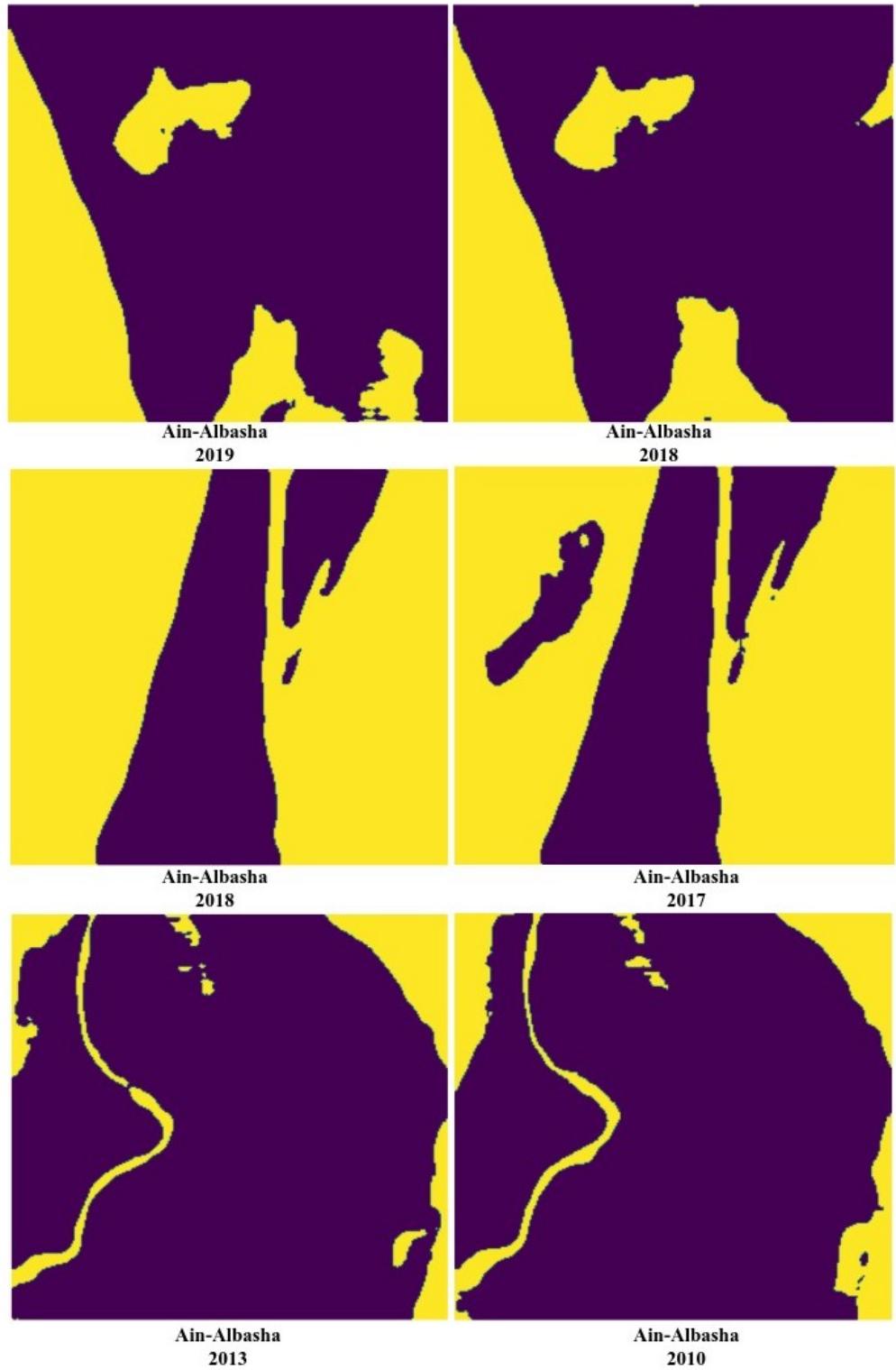


Figure 15: Smample of predicted images of Ain-Albasha region.

Chapter Five: Results and Discussions

This chapter presents and analyzes the experimental results of the proposed model. Section 5.1 illustrates the experiments configurations; Section 5.2 introduces the programming language and libraries; Section 5.3 presents the results of semantic segmentation; Section 5.4 analyses the performance evaluation.

5.1 Experiments Configuration

5.1.1 Environment Setup

All the experiments were carried out on the Jupyter Lab platform provided by Anaconda, which enables us to implement the entire project in Python. Table 8 summaries the details of the work environment include software and hardware that have been used in our work.

Table 8: Configuration of hardware and software

Operating System	Linux(Ubuntu 21.04)
GPU	Nvidia GEFORCE GTX 1660 Ti 6GB CUDA 11.2
CPU	Intel core i7
Ram	15.4 GiB
Platform	Anaconda (Jupyter)

5.1.2 Programming Language and Libraries

As described in chapter 4, we conduct thorough experiments on a set of predefined CNN architectures, the aim of the experiments carried out to figure out the best-performing ones on the forest dataset. Consequently, we examined the performance of three distinct CNNs under the same experimental configurations. These configurations are important as they largely affect the overall accuracy of any trained model.

There are several deep learning resources available to use for developers that made the implementation components and procedures more efficient. We used several toolkits and libraries after careful consideration and based on the unique requirements and time limitations of our project. We introduce them in the following subsections briefly.

5.1.2.1 Python for Data Science

Python is a high-level programming language that has various library and community capabilities for data science applications. We have used the following libraries:

- **NumPy:** is the basic package in python scientific computing. It provides a multidimensional array object, matrices, and various operations on them.
- **MatPlotLib:** a popular multi-functional 2D and 3D plotting library.
- **Panda:** provide high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

5.1.2.2 PyTorch

Torch is an open-source machine learning library, a scientific computing framework that provides a lot of deep learning algorithms.

- **nn:** is a package that used for building neural networks. It is divided into modular objects which have a forward() method to feed-forward and backward() method for feed backward.
- **util:** represents a Python iterable over a dataset.
- **Autograd:** performs the backpropagation starting from a variable.
- **Torchvision:** is a library for Computer Vision. It has utilities for efficient video and image transformations, some generally used pre-trained models.
- **Albumentations:** is a library for image augmentations. Albumentations supports different computer vision tasks such as semantic segmentation, instance segmentation, classification, instance segmentation, object detection, and pose estimation.

5.1.3 Hyper-parameters Optimization

Each input image (I) is resized to $256 \times 256 \times 3$. Imagenet uses mean and std as recommended. They have been calculated on millions of images. Calculating or using ImageNet's mean and std based on the user dataset. All CNN networks are trained with an initial learning rate (η) 0.001 using Adam optimizer (Opt). A batch size (Bs) of 10 is adopted in all

the experiments of model initialization and transfer learning. Each individual network is trained for 10 and 30 and 50 epochs (Te). We have conducted the model experiments using 30 epochs after experiments that showed stability of accuracy after epoch 30. This confirms the ability of our model in achieving fast learning and convergence. Table 9 summarizes these initialized hyper-parameters.

Table 9: Hyper parameters used in model initialization.

Training parameter	Value
Te	10,30,50
η	0.001
Opt	Adam
Bs	10
I	256 x 256 x 3

5.1.4 Evaluation Metrics

The following evaluation metrics used to measure the performance of the proposed model.

- **False Positive (FP):** it indicates an outcome where the model incorrectly predicts the positive class, i.e., forest.
- **False Negative (FN):** it indicates an outcome where the model incorrectly predicts the negative class, e., non-forest.
- **True Positive (TP):** it indicates an outcome where the model correctly predicts the positive class, i.e., forest.
- **True Negative (TN):** it indicates an outcome where the model correctly predicts the negative class, i.e., non-forest.
- **Mean Intersection-Over-Union (MIoU):** In the domain of image segmentation, the value of MIoU is an essential indicator to measure the performance of image segmentation methods. The IoU is calculated for each class at the pixel-level as:

$$IoU = \frac{TP}{TP + FP + FN} \quad (5.1)$$

Then the mean is taken for them.

- **Accuracy:** The accuracy of the module is calculating using the next formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.2)$$

- **Loss function:** Every pixel of the network’s output is compared with the corresponding pixel in the ground truth image. We employ standard cross-entropy loss on each pixel. The cross entropy is defined as:

$$E = \sum_{x \in \Omega} \omega(x) \log(p_{l(x)}(x)) \quad (5.3)$$

where $l : \Omega \rightarrow \{1, \dots, K\}$ the true label of each pixel (p), and $\omega : \Omega \rightarrow R$ is a weight map to give the pixels more interest in the training.

5.2 The Results of Semantic Segmentation

Obviously, the proposed architecture shows a noticeable impact on the accuracy results of U-net with CNN’s models. This initial finding is important to highlight how different convolutional features of forest images do act and perform on this dataset. Figures 16, 17 and 18 show that there is no large fluctuation in the training accuracy throughout the training and validation procedures. This also confirms that the proposed model overcomes any possible overfitting or underfitting issues.

In Figure 16, we notice that the Mobilenet_V2, in terms of training accuracy, is increasing gradually from 0.84 to reach 0.95, while the loss values are decreasing from 0.35 to reach 0.12. It can be obviously observed that the loss rate of the proposed model is smoothly reduced to lower values after epoch 5.

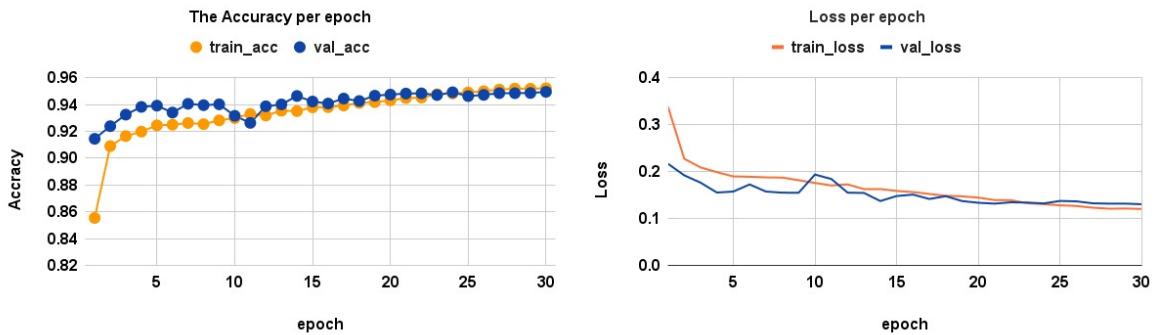


Figure 16: The accuracy and the loss of Mobilenet_V2 model.

In Figure 17, we notice that Mobilenet_V2, in terms of training accuracy, started from 0.88 and continued to increase with small percentages until it reached 95%. For the loss

rates, it gradually decreased from 0.3 until it reached 0.13. We can observe the consistency between the accuracy and loss.

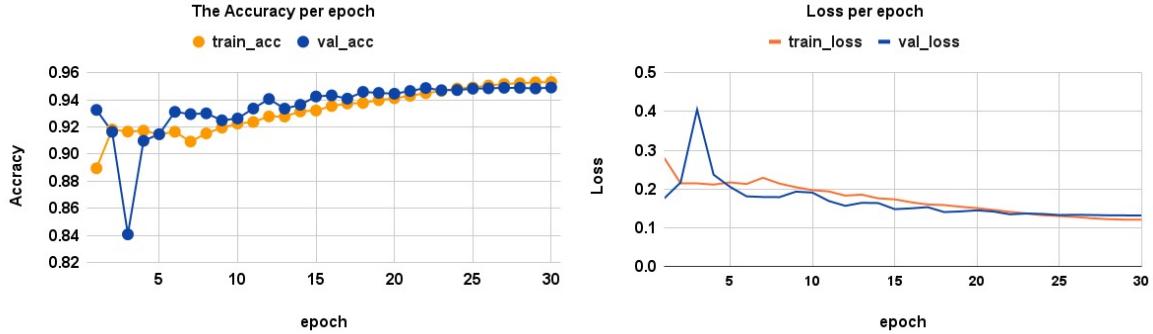


Figure 17: The accuracy and the loss of VGG16 model.

In Figure 18, the accuracy of the training began to increase from 0.87 to 0.91 in the first three epochs, then gradually increased by small percentages until it reached 0.94 at 30 epochs. Accordingly, the effect of the increase is reflected in the amount of the loss, as it started to decrease from 0.3 until it reached 0.15.

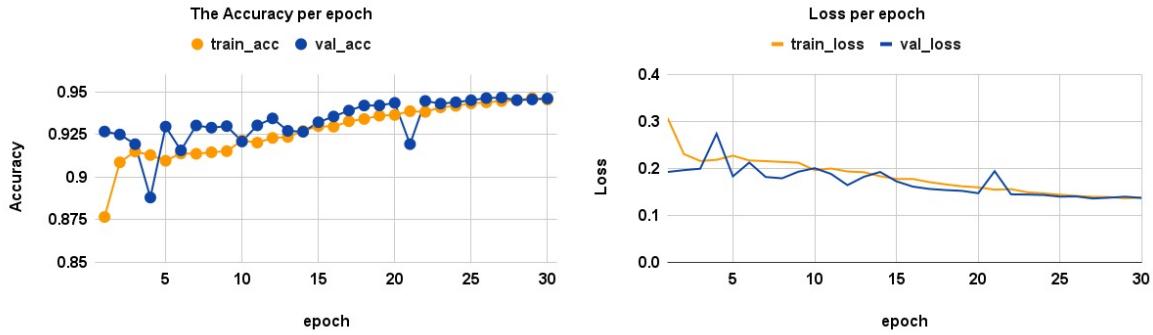


Figure 18: The accuracy and the loss of Resnet50 model.

Table 11 shows that Mobilenet_V2 also achieved the highest accuracy of 0.952, 0.951, and 0.948 for training, validation and testing, respectively. It was followed by VGG16 which achieved 0.953, 0.949 and 0.947 for training, verification and testing, respectively. VGG16 takes a shorter time than Mobilenet_V2 because it includes fewer number layers. This also confirms that the proposed model overcomes any possible overfitting or underfitting issues. The recorded accuracy indicates that our proposed model has performed very well.

Figure 19 shows the accuracy and the MIoU for each model, where we can notice that the accuracy rates are very close while the MIoU values of Mobilenet_V2 is greater than VGG16 and Resnet50.

Table 10: The accuracy of base models using 30 epochs.

Base Model	Train	Validation	Test	Time(hour)
Mobilenet_v2	0.952	0.951	0.948	20
VGG16	0.953	0.949	0.947	17
ResNet50	0.945	0.946	0.943	12

Table 11: The MIoU of base models using 30 epochs.

Base Model	Train	Validation	Test
Mobilenet_v2	0.875	0.861	0.821
VGG16	0.867	0.848	0.816
ResNet50	0.849	0.844	0.807

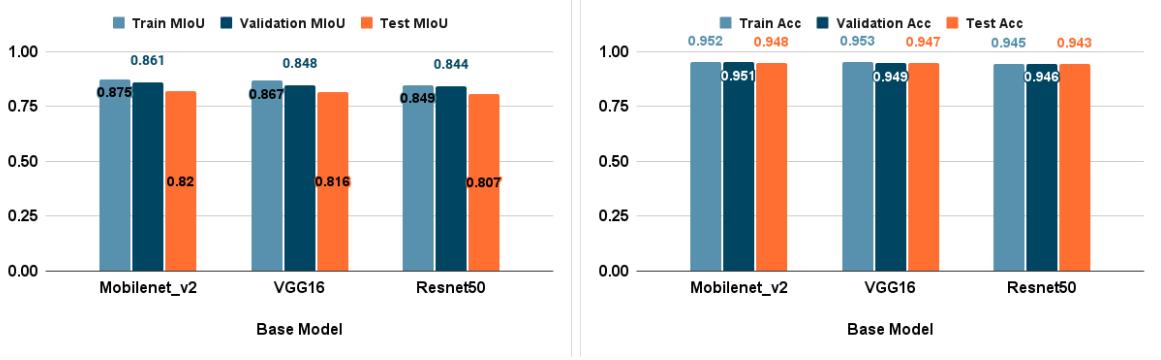


Figure 19: The MIoU and the accuracy of base models.

We cannot conduct a fair comparison with the related work using semantic segmentation for monitoring deforestation because we use our own collected images of Jordan forests. This bench-marking dataset includes different observed settings, areas, ground-truth images.

5.3 Deforestation Monitoring Performance

We monitored the change in forests over the years for a single area. We were able to track the change in 5 regions over different years and calculate the percentage of change that occurred in the area in terms of forest loss or gain.

Figure 20 shows examples of loss and gain results in forests for 3 random areas. The purple color indicates the forests and the yellow color indicates the non-forest. We calculated the loss and gain in the forest by calculating the amount of change in each color in the image. Figure (a) includes a small part of Ajloun forest area, in which the red square indicates

the area of change in forests. The percentage of the forest is 0.17, which is a very small percentage that is almost invisible, but the model was able to calculate and determine it. In figure (b) and (c), another example of the forest loss that occurs in the forests in the Yadouda area and Thaghret Asfour. The red squares indicate the areas of loss, where the loss occurred at a percentage of 1.17 and 0.26, respectively.



Figure 20: Sample of changes in various regions around the years.

Table 12 shows the percentage of change in the forests in different areas. We were able to measure only 5 regions because the images of the test dataset were chosen randomly. It can be noted in the table, each row in the first column represents two images of the same region in different years, and the second column gives the percentage of the difference between the two images. This variance ratio indicates the amount of loss or gains in the forest. According to the experimental results on deforestation monitoring, it can be seen that the loss of Zobia forests prevailed over the area. In contrast, in Thaghret Asfour area, the largest portion of the gain was greater than the loss. In Yadouda area, there are loss and gain areas in the same forest.

Table 12: The loss and gain rates over different years in 5 forest regions.

Zobia	difference Percentage	Loss/gain
img25_2019 & img30_2013	0.42%	Loss
img36_2016 & img39_2010	0.28%	Loss
img40_2020 & img45_2014	0.70%	Loss
img51_2017 & img55_2010	1.17%	Loss
img56_2020 & img57_2014	0.86%	Loss
img76_2014 & img78_2010	2.30%	Loss

Zayy	difference Percentage	Loss/gain
img21_2020 & img25_2014	0.42%	Loss
img34_2019 & img35_2018	1.17%	Gain

Thaghret-Asfour	difference Percentage	Loss/gain
img134_2019 & img135_2017	0.82%	Gain
img145_2014 & img147_2010	3.12%	Gain
img204_2019 & img208_2014	0.26%	Gain
img220_2016 & img223_2014	1.08%	Gain
img244_2014 & 246_2010	1.13%	Gain
img250_2013 & img252_2011	0.32%	Gain
img44_2017 & img46_2015	0.26%	Loss

Yadodah	difference Percentage	Loss/gain
img109_2012 & img106_2015	2.29%	Loss
img22_2019 & img23_2018	1.70%	Loss
img48_2013 & img50_2010	2.14%	Gain
img72_2019 & img79_2012	2.58%	Gain

Dair-Smadieh	difference Percentage	Loss/gain
img70_2019 & img72_2010	0.69%	Gain
img49_2019 & img51_2013	1.48%	Gain
img92_2015 & img93_2014	0.72%	Loss

Chapter Six: Conclusions and Future Work

6.1 Conclusions

In this thesis, we investigated various techniques and solutions for monitoring deforestation in Jordan forests using satellite imagery. The goal of our study is to develop a deep learning model for identifying deforestation based on semantic segmentation and discriminative features extracted from deep convolutional neural networks in order to achieve automated deforestation monitoring. We first studied the background of deforestation in Jordan and collected satellite imagery for the forests and green landscapes in Jordan. This dataset represented many regions over many years to monitor any changes in the forest coverage areas. Then, we created annotations for every single image. These ground truths only show the forest areas.

The proposed deep learning framework consists of three stages. The first stage presents the image pre-processing including techniques used such as data augmentation to classify the original dataset into two types, i.e. forest and non-forest. The second stage represents the model initialization by U-net and three pre-trained CNN models to extract a set of discriminating features for each forest image. The third stage represents feature extraction and semantic segmentation.

The proposed deep learning with U-Net and MobileNet_V2 achieves an accuracy of 0.948 and MIoU of 0.821. In terms of training time, it took about 20 hours which is more than the rest models. As a result, our model is able to effectively monitor deforestation.

Our model can also be used in predicting desertification over a period of years in one region because deforestation leads usually to soil damage and dryness, thus leading to the occurrence of desertification.

However, some limitations could affect our work for example the shadows of the trees that appeared in some years, the angle of the image that may change around the years. Additionally, a lack of some images in certain years for the same forest or region. There are areas where images are available for 10 consecutive years and some images are missing in the years between 2020 and 2010. The accuracy of the drawing is another challenging task

that may affect the segmentation performance since there is no tool to draw the borders with an accuracy of 100%. Also, considering some of the scattered trees as part of the forest or not is another important concern because they could be included or excluded from the forest.

6.2 Future Work

In future work, we plan to investigate the following:

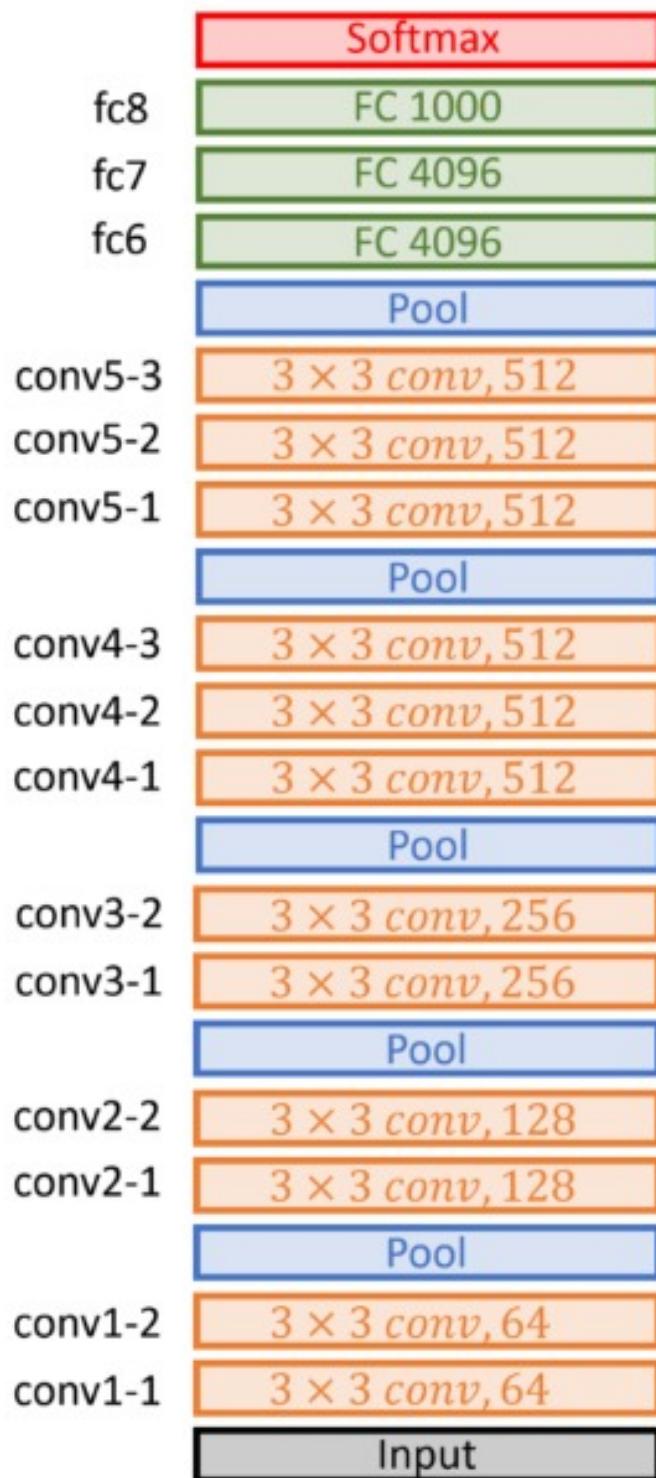
- Increasing the number of collected images either real-world satellite images or synthetic images by data augmentation.
- Considering more deep architectures,(e.g. Autoencoders, AlexNet), and features representation (e.g. graph representation).
- Attempting to find out more accurate approach to enhance the drawing accuracy of forest masks in the ground truth images.
- Someone can utilize the new collected images of Jordan forests as a benchmarking dataset for other applications such as forest topographic monitoring and forest analysis.

Appendix One: Specifications of Pretrained CNNs

Architectures

Input	Operator	<i>t</i>	<i>c</i>	<i>n</i>	<i>s</i>
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	-

Figure 21: MobileNetV2 blocks architecture [3].



VGG16

Figure 22: VGG16 architecture [4].
41

34-layer residual

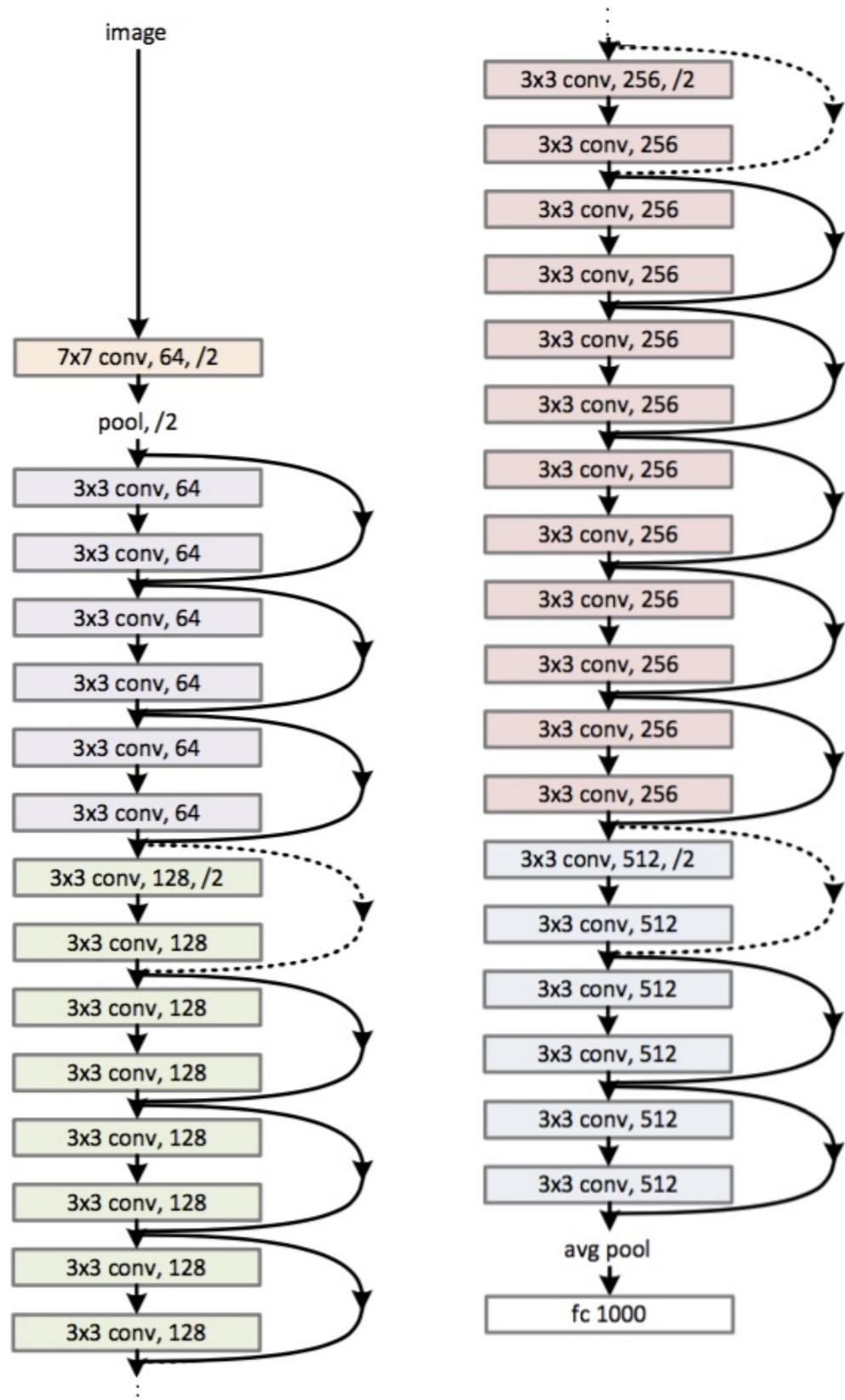


Figure 23: Resnet50 architecture [5].
42

References:

1. Remote Sensing for Forest Landscapes. | by Alexander Watson | openforests | Medium;. Available from: <https://medium.com/openforests/remote-sensing-for-forest-landscapes-83e246261c21>.
2. Deforestation and Forest Degradation | Threats | WWF;. Available from: <https://www.worldwildlife.org/threats/deforestation-and-forest-degradation>.
3. Hofman S. Federated Learning for Mobile and Embedded Systems [Ph.D. thesis]; 2020.
4. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. 2015.
5. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–778.
6. Maslin M, Malhi Y, Phillips O, Cowling S. New views on an old forest: assessing the longevity, resilience and future of the Amazon Rainforest. Transactions of the Institute of British Geographers. 2005 12;30:477 – 499.
7. Saatchi SS, Soares JV, Alves DS. Mapping deforestation and land use in amazon rainforest by using SIR-C imagery. Remote Sensing of Environment. 1997;59(2):191–202. Spaceborne Imaging Radar Mission.
8. Müller C. Brazil and the Amazon Rainforest: Deforestation, Biodiversity and Cooperation with the EU and International Forums. European Parliament; 2020.
9. Eccleston P. Half of Amazon could be gone by 2030. Telegraph 6Dec <http://www.telegraph.co.uk/earth/main.jhtml>. 2007.
10. Isaienkov K, Yushchuk M, Khramtsov V, Seliverstov O. Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem With Sentinel-2. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020;14:364–376.
11. Mikhaylov A, Moiseev N, Aleshin K, Burkhardt T. Global climate change and greenhouse effect. Entrepreneurship and Sustainability Issues. 2020;7(4):2897.
12. Planet | Homepage;. Available from: <https://www.planet.com/>.
13. Forest Monitoring, Land Use & Deforestation Trends | Global Forest Watch;. Available from: <https://www.globalforestwatch.org/>.

14. Zhang L, Zhang L, Du B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*. 2016;4(2):22–40.
15. Scott GJ, England MR, Starms WA, Marcum RA, Davis CH. Training deep convolutional neural networks for land-cover classification of high-resolution imagery. *IEEE Geoscience and Remote Sensing Letters*. 2017;14(4):549–553.
16. Chen F, Ren R, Van de Voorde T, Xu W, Zhou G, Zhou Y. Fast automatic airport detection in remote sensing images using convolutional neural networks. *Remote Sensing*. 2018;10(3):443.
17. Kemker R, Salvaggio C, Kanan C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS journal of photogrammetry and remote sensing*. 2018;145:60–77.
18. Remote Sensing for Forest Landscapes. | by Alexander Watson | openforests | Medium;. Available from: <https://medium.com/openforests/remote-sensing-for-forest-landscapes-83e246261c21>.
19. What Is Deforestation? Definition, Causes, Effects and Solutions to Stop it;. Available from: <https://youmatter.world/en/definition/definitions-what-is-definition-deforestation-causes-effects/>.
20. Campbell JB, Wynne RH. Introduction to remote sensing. Guilford Press; 2011.
21. Davis SM, Landgrebe DA, Phillips TL, Swain PH, Hoffer RM, Lindenlaub JC, et al. Remote sensing: the quantitative approach. New York. 1978.
22. Netzband M, Stefanov WL, Redman C. Applied remote sensing for urban planning, governance and sustainability. Springer Science & Business Media; 2007.
23. Johnson EA, Miyanishi K. Forest fires. Behaviour and ecological effects Elsevier. 2001;570.
24. Flannigan MD, Stocks BJ, Wotton BM. Climate change and forest fires. *Science of the total environment*. 2000;262(3):221–229.
25. Ouahabi A, Taleb-Ahmed A. Deep learning for real-time semantic segmentation: Application in ultrasound imaging. *Pattern Recognition Letters*. 2021;144:27–34.
26. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*. 2016;316(22):2402–2410.

27. Jia X, Meng MQH. A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2016. p. 639–642.
28. Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*. 1997;8(1):98–113.
29. Mhatre A, Mudaliar NK, Narayanan M, Gurav A, Nair A, Nair A. Using Deep Learning on Satellite Images to Identify Deforestation/Afforestation. In: International Conference On Computational Vision and Bio Inspired Computing. Springer; 2019. p. 1078–1084.
30. Rakshit S, Debnath S, Mondal D. Identifying land patterns from satellite imagery in amazon rainforest using deep learning. *arXiv preprint arXiv:180900340*. 2018.
31. Bragagnolo L, da Silva RV, Grzybowski JMV. Amazon forest cover change mapping based on semantic segmentation by U-Nets. *Ecological Informatics*. 2021;62:101279.
32. Chantharaj S, Pornratthanapong K, Chitsinpchayakun P, Panboonyuen T, Vateekul P, Lawavirojwong S, et al. Semantic segmentation on medium-resolution satellite images using deep convolutional networks with remote sensing derived indices. In: 2018 15th International joint conference on computer science and software engineering (JCSSE). IEEE; 2018. p. 1–6.
33. Brovelli MA, Sun Y, Yordanov V. Monitoring forest change in the amazon using multi-temporal remote sensing data and machine learning classification on Google Earth Engine. *ISPRS International Journal of Geo-Information*. 2020;9(10):580.
34. Wyniawskyj NS, Napiorkowska M, Petit D, Podder P, Marti P. Forest monitoring in guatemala using satellite imagery and deep learning. In: IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE; 2019. p. 6598–6601.
35. Ortega M, Bermudez J, Happ P, Gomes A, Feitosa R. EVALUATION OF DEEP LEARNING TECHNIQUES FOR DEFORESTATION DETECTION IN THE AMAZON FOREST. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*. 2019;4.
36. Lee SH, Han KJ, Lee K, Lee KJ, Oh KY, Lee MJ. Classification of Landscape Affected by Deforestation Using High-Resolution Remote Sensing Data and Deep-Learning Techniques. *Remote Sensing*. 2020;12(20):3372.
37. Maretto RV, Fonseca LM, Jacobs N, Körting TS, Bendini HN, Parente LL. Spatio-temporal deep learning approach to map deforestation in amazon rainforest. *IEEE Geoscience and Remote Sensing Letters*. 2020;18(5):771–775.

38. Ban Y, Zhang P, Nascetti A, Bevington AR, Wulder MA. Near real-time wildfire progression monitoring with Sentinel-1 SAR time series and deep learning. *Scientific Reports*. 2020;10(1):1–15.
39. Andrade R, Costa G, Mota G, Ortega M, Feitosa R, Soto P, et al. Evaluation of semantic segmentation methods for deforestation detection in the amazon. *ISPRS Archives*; volume 43, B3. 2020;43(B3):1497–1505.
40. Taghanaki SA, Abhishek K, Cohen JP, Cohen-Adad J, Hamarneh G. Deep semantic segmentation of natural and medical images: a review. *Artificial Intelligence Review*. 2021;54(1):137–178.
41. Manisha P, Jayadevan R, Sheeba VS. Content-based image retrieval through semantic image segmentation. In: *AIP Conference Proceedings*. vol. 2222. AIP Publishing LLC; 2020. p. 030008.
42. Ouni A, Royer E, Chevaldonné M, Dhome M. Leveraging semantic segmentation for hybrid image retrieval methods. *Neural Computing and Applications*. 2021:1–19.
43. Pohlen T, Hermans A, Mathias M, Leibe B. Full-resolution residual networks for semantic segmentation in street scenes. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2017. p. 4151–4160.
44. Zulfiqar A, Ghaffar MM, Shahzad M, Weis C, Malik MI, Shafait F, et al. AI-ForestWatch: semantic segmentation based end-to-end framework for forest estimation and change detection using multi-spectral remote sensing imagery. *Journal of Applied Remote Sensing*. 2021;15(2):024518.
45. da Costa LB, de Carvalho OLF, de Albuquerque AO, Gomes RAT, Guimarães RF, de Carvalho Júnior OA. Deep Semantic Segmentation for Detecting Eucalyptus Planted Forests in the Brazilian Territory Using Sentinel-2 Imagery. *Geocarto International*. 2021;(just-accepted):1–12.
46. Wang L, Zhou Y, Hu Q, Tang Z, Ge Y, Smith A, et al. Early Detection of Encroaching Woody Juniperus virginiana and Its Classification in Multi-Species Forest Using UAS Imagery and Semantic Segmentation Algorithms. *Remote Sensing*. 2021;13(10):1975.
47. Zhang B, Mu H, Gao M, Ni H, Chen J, Yang H, et al. A Novel Multi-Scale Attention PFE-UNet for Forest Image Segmentation. *Forests*. 2021;12(7):937.
48. Garg R, Kumar A, Bansal N, Prateek M, Kumar S. Semantic segmentation of PolSAR image data using advanced deep learning model. *Scientific reports*. 2021;11(1):1–18.
49. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *Journal of Big Data*. 2019;6(1):1–48.

50. Wang J, Perez L, et al. The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks Vis Recognit.* 2017;11:1–8.
51. Lashgari E, Liang D, Maoz U. Data augmentation for deep-learning-based electroencephalography. *Journal of Neuroscience Methods.* 2020;108885.
52. Wagner FH, Sanchez A, Tarabalka Y, Lotte RG, Ferreira MP, Aidar MP, et al. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation.* 2019;5(4):360–375.
53. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:14091556.* 2014.
54. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition;* 2016. p. 770–778.
55. Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:170404861.* 2017.
56. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition;* 2018. p. 4510–4520.
57. Amjoud AB, Amrouch M. Convolutional neural networks backbones for object detection. In: *International Conference on Image and Signal Processing.* Springer; 2020. p. 282–289.
58. Arnab A, Zheng S, Jayasumana S, Romera-Paredes B, Larsson M, Kirillov A, et al. Conditional random fields meet deep neural networks for semantic segmentation: Combining probabilistic graphical models with deep learning for structured prediction. *IEEE Signal Processing Magazine.* 2018;35(1):37–52.
59. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition;* 2015. p. 3431–3440.

مراقبة انحسار الغابات في الأردن باستخدام التجزئة الدلالية والتعلم العميق

إعداد: لوجين حامد محمد الصمادي

الملخص

يشهد الأردن تحولات ضخمة في طبيعة البيئة و السمات الطبوغرافية ، حيث تمثل الصحراe جزءاً كبيراً من أراضي الأردن مع مساحة غابات محدودة للغاية. على مدى العقود الثلاثة الماضية ، فقد الأردن قرابة ثلث الغابات الطبيعية بمعدل 1.6 % سنوياً.

للمساعدة في مراقبة إزالة الغابات في الأردن ، قمنا بجمع مجموعة بيانات كبيرة من غابات الأردن في شكل صور عالية الدقة من الأقمار الصناعية. الهدف الرئيسي من هذه الأطروحة هو تطوير نموذج التعلم العميق لمراقبة إزالة الغابات تلقائياً بناءً على التجزئة الدلالية. جدير بالذكر أن هناك نقصاً في الدراسات التي تعمل على مراقبة إزالة الغابات باستخدام التجزئة الدلالية مع الشبكات العصبية العميقه. لذلك ، نساهم في جمع مجموعة بيانات جديدة من غابات الأردن على مدى عشر سنوات ، بالإضافة إلى توفير حل دقيق يدعم الذكاء الاصطناعي لمراقبة إزالة الغابات في الأردن. يشتمل النموذج المقترن على أربع مراحل أساسية: المعالجة المساعدة للبيانات ، وتهيئة البنية العميقه ، واستخراج الميزات ، والتجزئة الدلالية. استخدمنا بنية U-Net الفعالة ، أي نموذج التشفير / وحدة فك التشفير ، لاستخراج مجموعة من الميزات التمييزية للتجزئة الدلالية من أجل التنبؤ بإزالة الغابات في صور الاختبار. يتم أيضاً إعداد صور الحقيقة المطابقة مع مراعاة متأنية لأقتعة الغابة المرسومة لكل غابة.

بالإضافة إلى ذلك ، قمنا بتهيئة وفحص التصميم الخاص بنا من خلال أحد النماذج الثلاثة المدربة مسبقاً: Resnet50 و VGG-16 و VGG16 Resnet50 و MobilenetV2. يتم ضبط هذه النماذج العميقه المستندة إلى CNN بدقة في مجال صور الغابة من خلال إجراء نقل التعلم من مجال الصور ذات الأغراض العامة. أظهر أداء عملية التجزئة الدلالية قدرة عالية في اكتشاف الغابة والتنبؤ بالمناطق المحيطة. علاوة على ذلك ، أثبتت النتائج التجريبية فعالية النموذج في التنبؤ بمعدلات الخسارة والكسب في الغابات من خلال الإبلاغ عن دقة تبلغ 94.8 % ، و MIoU بنسبة 82 %. كما أظهر النموذج المقترن دقة مماثلة باستخدام Mobilenet_V2. يتحقق Resnet50 و VGG16 على التوالي. نجح النموذج المقترن في تحديد الغابة المتأثرة بأي نوع من إزالة الغابات على مدى مجموعة من السنوات ، وهو قادر على تقدير النسبة المئوية لتغير الغابات ، أي الربح أو الخسارة ، بعد تطبيق فحص التشابه على الغابات قيد الدراسة.