REPUBLIC OF TÜRKİYE
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

BURNED AREA SEGMENTATION AND SEVERITY ESTIMATION IN POST FIRE LANDSCAPES USING DEEP LEARNING METHODS

Murat Mert YURDAKUL

MASTER OF SCIENCE THESIS
Department of Computer Engineering
Program of Computer Engineering

Supervisor
Assoc. Prof. Dr. Hamza Osman İLHAN

March, 2025

REPUBLIC OF TÜRKİYE
YILDIZ TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

BURNED AREA SEGMENTATION AND SEVERITY ESTIMATION IN POST FIRE LANDSCAPES USING DEEP LEARNING METHODS

A thesis submitted by Murat Mert YURDAKUL in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE** is approved by the committee on 25.03.2025 in Department of Computer Engineering, Program of Computer Engineering.

Assoc. Prof. Dr. Hamza Osman İLHAN
Yildiz Technical University
Supervisor

Approved By the Examining Committee

Assoc. Prof. Dr. Hamza Osman İLHAN, Supervisor
Yildiz Technical University _____

Assoc. Prof. Dr. Ali Can KARACA, Member
Yildiz Technical University _____

Asst. Prof. Dr. Ali SARIKAŞ, Member
Marmara University _____

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of BURNED AREA SEGMENTATION AND SEVERITY ESTIMATION IN POST FIRE LANDSCAPES USING DEEP LEARNING METHODS supervised by my supervisor, Assoc. Prof. Dr. Hamza Osman İLHAN. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Murat Mert YURDAKUL

Signature

This study was supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) ARDEB 2549 (NCBR) Grant No: 122N254.

Dedicated to my family
and my friends

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my family for their unwavering support, patience, and encouragement throughout this journey. Their belief in me has been a constant source of strength, making this challenging endeavor more manageable. I am also sincerely grateful to my supervisor, Assoc. Prof. Dr. Hamza Osman İLHAN, for his invaluable guidance, constructive feedback, and continuous encouragement. His expertise has been instrumental in shaping this research, and I truly appreciate his mentorship. Lastly, I extend my appreciation to everyone who has contributed to this work, directly or indirectly, for their assistance, discussions, and motivation.

Murat Mert YURDAKUL

TABLE OF CONTENTS

LIST OF SYMBOLS viii

LIST OF ABBREVIATIONS ix

LIST OF FIGURES xi

LIST OF TABLES xii

ABSTRACT xiii

ÖZET xv

1 INTRODUCTION 1
- 1.1 Goal And Objectives . . . 1
- 1.2 Motivation and Contributions . . . 2
- 1.3 Research Methodology . . . 2
- 1.4 Overview . . . 2

2 LITERATURE REVIEW 4

3 UNDERSTANDING WILDFIRES 10
- 3.1 Definition of Wildfires . . . 10
- 3.2 Causes of Wildfires . . . 10
- 3.3 Impacts of Wildfires . . . 11
- 3.4 Wildfire Trends . . . 12
- 3.5 Post-Fire Land Management Strategies and Severity Assessment . . . 14

4 MATERIAL AND METHOD 17
- 4.1 Dataset . . . 17
- 4.2 Double-Step Deep Learning Framework [12] . . . 19
  - 4.2.1 Step 1: Wildfire Detection . . . 19
  - 4.2.2 Step 2: Severity Estimation . . . 20
  - 4.2.3 Implemented Models . . . 21
- 4.3 Proposed Methods . . . 26
  - 4.3.1 Mask R-CNN . . . 27

4.3.2 MultiResUNet 28
4.3.3 Level 4 and Level 5 Experiments 31
4.3.4 Hyperparameters 32
4.3.5 Performance Metrics 34

5 EXPERIMENTAL RESULTS **36**
5.1 Mask R-CNN 36
5.2 DSF with MultiResUNet 37
5.2.1 Level 4 and 5 Experiments 38
5.2.2 Combined Pre- and Post-Fire Imagery 41
5.2.3 Decoder Batch Normalization Removed 41

6 CONCLUSION **44**
6.1 Future Works 46

REFERENCES **48**

PUBLICATIONS FROM THE THESIS **52**

LIST OF SYMBOLS

y_i	Actual Severity Level for a Given Burned Area (RMSE)
α	Attention Coefficients (Attention U-Net)
CO_2	Carbon Dioxide
X	Convolution (Nested U-Net)
D_g	Depth of the Decoder Feature Map (Attention U-Net)
D_x	Depth of the Encoder Feature Map (Attention U-Net)
ha	Hectare
H_g	Height of the Decoder Feature Map (Attention U-Net)
H_x	Height of the Encoder Feature Map (Attention U-Net)
g	High-level Feature Map from the Decoder path (Attention U-Net)
X^l	Low-level Feature Map from the Encoder Path (Attention U-Net)
$\hat{X}^l$	Refined Feature Map (Attention U-Net)
CH_4	Methane
$\hat{y_i}$	Predicted Severity Level (RMSE)
$\mathcal{L}$	Refinement Levels (Nested U-Net)
σ_1	ReLU Activation Function (Attention U-Net)
σ_2	Sigmoid Activation Function (Attention U-Net)
N	Total Number of Observations (RMSE)
ψ	Weights of the Convolution Applied after the ReLU Activation (Attention U-Net)
W_g	Width of the Decoder Feature Map (Attention U-Net)
W_x	Width of the Encoder Feature Map (Attention U-Net)

LIST OF ABBREVIATIONS

AOD	Aerosol Optical Depth
BAER	Burned Area Emergency Response
BCE	Binary Cross-Entropy
DSF	Double-Step (Deep Learning) Framework
EEA	European Environment Agency
EMS	Copernicus Emergency Management Service
FARSITE	Fire Area Simulator
FN	False Negative
FP	False Positive
GNDVI	Green Normalized Difference Vegetation Index
IoU	Intersection over Union
JRC	Joint Research Center
LSTM	Long-Short-Term-Memory
MSE	Mean Square Error
MSI	Multispectral Instrument
MTL	Multi-Task Learning
NASA	National Aeronautics and Space Administration
NAIP	National Agriculture Imagery Program
NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation IndeX
NIFC	National Interagency Fire Center
NIR	Near-Infrared
OGM	Orman Genel Müdürlüğü

PM2.5	Fine Particulate Matter
R-CNN	Region-Based Convolutional Neural Network
RGB	Red-Green-Blue
RMSE	Root Mean Square Error
RPN	Region Proposal Network
SAR	Synthetic Aperture Radar
SBS	Soil Burn Severity
SWIR	Shortwave Infrared
TP	True Positive
VGG	Visual Geometry Group
VH	Vertical-Horizontal
VIIRS	Visible Infrared Imaging Radiometer Suite
VV	Vertical-Vertical
WHP	Wildfire Hazard Potential

LIST OF FIGURES

Figure 3.1 Total output according to cause of fire, 2023 [20] 11
Figure 3.2 US wildfires since 2020 [23] 13
Figure 4.1 Logarithmic representation of pixels from mask images based on folds 18
Figure 4.2 Double-Step framework architecture [12] 19
Figure 4.3 U-Net architecture [27] 21
Figure 4.4 Attention U-Net architecture[28] 23
Figure 4.5 Attention gate[28] 24
Figure 4.6 Nested U-Net architecture[29] 25
Figure 4.7 Mask R-CNN architecture[31] 27
Figure 4.8 MultiResUNet architecture[13] 28
Figure 4.9 MultiRes block architecture[13] 29
Figure 4.10 Res path architecture[13] 30
Figure 5.1 Example of predicted images 38
Figure 5.2 Example of MultiResUNet predictions based on levels 40

LIST OF TABLES

Table 3.1 Wildfire Statistics in Türkiye (2020-2023) [20] 13
Table 4.1 The organization schema of the validation and test sets in the experiments 17
Table 4.2 Hyperparameter settings 33
Table 5.1 Step 1 (Wildfire detection phase) results - binary IoU 37
Table 5.2 Step 2 (Wildfire severity classification) results - regression RMSE 39
Table 5.3 Performance benchmark of related works 40
Table 5.4 MultiResUNet results with architectural variations 43

ABSTRACT

BURNED AREA SEGMENTATION AND SEVERITY ESTIMATION IN POST FIRE LANDSCAPES USING DEEP LEARNING METHODS

Murat Mert YURDAKUL

Department of Computer Engineering
Master of Science Thesis

Supervisor: Assoc. Prof. Dr. Hamza Osman İLHAN

As wildfires become more frequent and intense, it is essential to develop sophisticated techniques for precise detection and damage evaluation. Accurate detection and severity assessment of burned areas are critical for effective wildfire management. This study evaluates a Double-Step Deep Learning Framework designed to identify burned areas and classify their severity using high-resolution Sentinel-2 post-fire imagery. The proposed framework comprises two steps: a binary segmentation step to detect burned areas and a severity classification step that assigns levels of damage, ranging from unburned to completely destroyed. The study explores the effect of different severity levels within mask output, focusing on both 4 and 5 level severity classifications. Additionally, the Mask R-CNN model was evaluated independently for image segmentation, revealing challenges due to its reliance on pretrained weights and limited spectral input. Several U-Net-based architectures, including MultiResUNet, were implemented and evaluated within this framework. The results indicate that MultiResUNet achieved competitive performance in burned area detection but did not consistently outperform other architectures, such as Nested U-Net and Attention U-Net. Additionally, reducing severity levels from five to four improved model performance by addressing data imbalance issues. These findings highlight the strengths and limitations of the Double-Step Deep Learning Framework in advancing post-fire damage assessment and decision-making for wildfire management. While MultiResUNet provided

robust results, alternative U-Net-based architectures showed potential advantages in specific tasks, particularly in severity classification. This approach has the potential to improve the precision of damage assessments and support more informed decision-making in the management and response of wildfires.

Keywords: Wildfire severity prediction, deep neural networks, double-step architecture, satellite imagery analysis.

ÖZET

DERİN ÖĞRENME YÖNTEMLERİ İLE YANGIN SONRASI ALANLARDA YANIK BÖLGE SEGMENTASYONU VE ŞİDDET TAHMİNİ

Murat Mert YURDAKUL

Bilgisayar Mühendisliği Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç. Dr. Hamza Osman İLHAN

Orman yangınlarının daha sık ve yoğun hale gelmesiyle birlikte, doğru tespit ve hasar değerlendirmesi için gelişmiş tekniklerin geliştirilmesi büyük önem taşımaktadır. Yanmış alanların doğru bir şekilde tespiti ve şiddet değerlendirmesi, etkili yangın yönetimi için kritik öneme sahiptir. Bu araştırma, yanmış alanları belirlemek ve yangın şiddetini tahmin etmek için yüksek çözünürlüklü Sentinel-2 yangın sonrası görüntüleri kullanan Çift Aşamalı Derin Öğrenme Çerçevesi'ni incelemektedir. Önerilen çerçeve iki aşamadan oluşmaktadır: yanmış alanları tespit etmek için ikili segmentasyon adımı ve hasar seviyelerini yanmamıştan tamamen yok edilmişe kadar atayan şiddet sınıflandırma adımı. Çalışma, maske çıktılarındaki farklı şiddet seviyelerinin etkisini incelemek için hem 4 hem de 5 seviyeli şiddet sınıflandırmalarına odaklanmaktadır. Ayrıca, Mask R-CNN modeli görüntü segmentasyonu için bağımsız olarak değerlendirilmiştir. Mask R-CNN modeli ile yapılan deneyde, önceden eğitilmiş ağırlıklara ve sınırlı spektral girdiye bağımlılık nedeniyle karşılaşılan zorluklar ele alınmıştır. MultiResUNet dahil olmak üzere birkaç U-Net tabanlı mimari bu çerçeve içinde uygulanmış ve değerlendirilmiştir. Sonuçlar, MultiResUNet'in yanmış alan tespitinde tatmin edici performans gösterdiğini ancak Nested U-Net ve Attention U-Net gibi diğer mimarileri tutarlı bir şekilde geride bırakamadığını göstermektedir. Ayrıca şiddet seviyelerinin beşten dörde düşürülmesi, veri dengesizliği sorunlarını çözerek model performansını artırmıştır. Bu bulgular, Çift Aşamalı Derin Öğrenme Çerçevesi'nin yangın sonrası hasar değerlendirmesini ve yangın yönetimi için karar

alma süreçlerini geliştirme potansiyelini vurgulamaktadır. MultiResUNet güçlü sonuçlar sağlasa da, özellikle şiddet sınıflandırması gibi belirli görevlerde alternatif U-Net tabanlı mimariler potansiyel avantajlar göstermiştir. Bu yaklaşım, hasar değerlendirmelerinin doğruluğunu artırma ve orman yangınlarının yönetimi ve müdahalesinde daha bilinçli kararlar alınmasını destekleme potansiyeline sahiptir.

Anahtar Kelimeler: Yangın şiddeti tahmini, derin sinir ağları, çift adımlı mimari, uydu görüntü analizi.

1
INTRODUCTION

The growing threat posed by wildfires in recent years has increased the need for advanced technologies to assess fire damage accurately and efficiently. The destruction caused by these fires affects ecosystems, economies and communities worldwide, making it essential to have reliable tools for fire detection and damage assessment. High-resolution satellite imagery from missions such as Sentinel-2 offers a critical resource for monitoring large-scale environmental events such as wildfires.

1.1 Goal And Objectives

The primary aim of this thesis is to investigate and compare the performance of the U-Net model and its variants with a newly proposed model, MultiResUNet, for wildfire damage assessment. This comparison explores their effectiveness in binary segmentation and severity regression tasks.

The objectives of the thesis can be summarized as follows:

- Model Comparison: To evaluate the similarities and differences between U-Net variants and MultiResUNet, focusing on their performance in binary segmentation and severity regression tasks.
- Severity Mask Levels: To analyze the impact of reducing severity mask levels on model performance, balancing improvements in accuracy with the potential loss of detail.
- Baseline Analysis: To assess the performance of models like Mask R-CNN when applied to the same dataset using a simplified setup with limited channels, identifying limitations and areas for improvement.

1.2 Motivation and Contributions

The motivation behind this research stems from the urgent need for effective tools for wildfire damage assessment. This thesis introduces and evaluates a novel approach, contributing to the field by:

- Proposing MultiResUNet as an enhanced architecture for satellite image analysis.
- Demonstrating the significance of severity mask optimization in improving framework performance.
- Highlighting the limitations of existing models under constrained conditions, providing insights for future development.

1.3 Research Methodology

This study adopts a systematic methodology comprising the following steps:

Dataset Preparation: Utilizing high-resolution satellite imagery from Sentinel-2 with preprocessing steps to ensure compatibility with different model architectures.

Model Training and Evaluation: Training U-Net variants and MultiResUNet on the same dataset, followed by a comparative analysis of their performance in binary segmentation and severity regression tasks.

Parameter Optimization: Experimenting with severity mask levels to evaluate their impact on model accuracy and efficiency.

Baseline Comparisons: Assessing the performance of baseline models like Mask R-CNN under simplified conditions.

1.4 Overview

The thesis is organized as follows:

1. Chapter 1 presents an overview of wildfire. This study's goal and research aims were also explained.
2. Chapter 2 reviews the existing literature on forest fire datasets, wildfire detection and damage assessment.

3. Chapter 3 explores the causes and impacts of wildfires, discussing environmental, economic, and social consequences while establishing the foundation for severity assessment.

4. Chapter 4 presents the methodology adopted for model training and evaluation.

5. Chapter 5 discusses the experimental results, comparing the performance of U-Net variants and MultiResUNet. Four-level and five-level severity mask results are analyzed. For MultiResUNet, the effects of removing decoder batch normalization and using a 24-channel input are also evaluated. Additionally, Mask R-CNN results are briefly discussed.

6. Chapter 6 provides conclusions and outlines potential directions for future research.

2
LITERATURE REVIEW

Wildfire segmentation and severity assessment rely on satellite datasets such as Sentinel-2, Landsat, and MODIS, which provide multispectral imagery for detecting burned areas. Sentinel-2, with its 12 spectral bands, is widely used for severity mapping due to its ability to capture post-fire spectral variations.

Colomba et al. introduced a new data set specifically designed to delineate the burned area and estimate severity from satellite imagery [1]. The authors used Sentinel-1 and Sentinel-2 satellite data to categorize burn severity on a scale from 0 (undamaged) to 4 (completely destroyed).

Navarro et al. conducted an analysis of 2016 forest fires on Madeira Island using Sentinel-2A imagery to evaluate the severity of the burn [2]. They focused on the application of spectral indices such as the Normalized Difference Vegetation Index (NDVI), Green Normalized Difference Vegetation Index (GNDVI) and Normalized Burn Ratio (NBR). By comparing these indices with data from the Copernicus Emergency Management Service, the authors confirmed that red-edge spectral indices were highly effective in identifying levels of burn severity.

The CaBuAr dataset [3], introduced by Cambrin, Colomba, and Garza, is an open-access resource for burned area delineation using Sentinel-2 satellite imagery. It includes 340 wildfire events from California (2015–2022), with pre-fire and post-fire Level-2A images, ground truth masks derived from California Department of Forestry and Fire Protection data, and 12 spectral bands at 20m resolution. The dataset enables binary segmentation, change detection, and spectral index evaluation, supporting deep learning models like U-Net and SegFormer, with U-Net achieving the highest F1-score. Preprocessed to remove cloud-covered or incomplete samples, it serves as a benchmark for wildfire segmentation, disaster recovery, and environmental monitoring. Future plans include geographic expansion, multimodal data integration, and resampled versions for various resolutions to enhance wildfire management and response.

The FLAME (Fire Luminosity Airborne-based Machine Learning Evaluation) dataset [4] is a comprehensive aerial fire detection and segmentation resource collected using drones during prescribed burns in Northern Arizona. It includes normal and thermal imagery in multiple spectral palettes, such as Fusion, WhiteHot, and GreenHot, capturing diverse fire conditions. The dataset consists of 39,000+ frames for classification and 2,000+ manually labeled fire masks for segmentation, spanning multiple resolutions and formats to support various fire monitoring applications. Designed to advance wildfire detection and management, FLAME facilitates research in fire behavior analysis, disaster response, and resource allocation. Future developments aim to integrate environmental data like terrain models and vegetation indices for improved fire spread modeling and risk assessment.

The FireRisk dataset [5] is an open-access resource designed for fire risk assessment using remote sensing imagery. It consists of 91,872 labeled images, categorized into seven fire risk levels, from "very low" to "very high", based on data from the Wildfire Hazard Potential (WHP) raster dataset and high-resolution imagery from the National Agriculture Imagery Program (NAIP). Unlike traditional fire risk models that rely on geospatial and environmental data, FireRisk solely uses RGB remote sensing images, each covering a 270x270 meter area, to classify fire-prone regions. The dataset was collected between 2019 and 2020, ensuring high-quality images with minimal cloud cover. Designed as a benchmark for machine learning-based fire risk assessment, FireRisk provides a simpler alternative to traditional geoscientific models. Future expansions aim to broaden geographic coverage and incorporate thermal imagery to enhance fire risk prediction and disaster prevention efforts.

The Wildfire Dataset [6] is designed to improve deep learning-based forest fire detection by addressing the high false alarm rates in existing systems. It consists of 2,700 aerial and ground-based RGB images, categorized into five subclasses that capture variations in fire, smoke, and confounding elements resembling fire or smoke. By incorporating diverse environmental conditions, forest types, and geographical regions, the dataset enhances model robustness in real-world scenarios. To improve detection accuracy, the authors propose a multi-task learning (MTL) framework, enabling simultaneous binary classification (fire/no fire) and multi-class classification (confounding elements), reducing false positives and enhancing overall model performance. The dataset is publicly available, accompanied by metadata and download links, addressing the scarcity of open-source forest fire datasets and providing a benchmark for future wildfire detection research.

The EO4WildFires [7] dataset is a benchmark resource designed to support wildfire severity prediction using multisensor Earth observation and meteorological data. It integrates Sentinel-2 multispectral imagery, Sentinel-1 SAR data, and meteorological parameters from NASA Power, covering 31,730 wildfire events across 45 countries (2018–2022), with annotations from the European Forest Fire Information System (EFFIS). Each wildfire event is structured as a data cube, capturing pre-fire environmental conditions, including vegetation indices, canopy structure, and climate variables such as temperature, precipitation, and soil moisture. The dataset enables machine learning models to predict wildfire severity, with experiments demonstrating that Sentinel-2 data is the most effective, followed by Sentinel-1 SAR, while meteorological data alone had lower predictive performance. EO4WildFires is open-access, aiming to advance early warning systems, resource allocation, and wildfire impact mitigation, with future expansions planned to incorporate additional data sources and advanced predictive models.

The Active Fire Detection in Landsat-8 Imagery study introduces a large-scale dataset designed for active fire detection using Landsat-8 satellite imagery [8]. The dataset consists of over 150,000 image patches with 10-band spectral data, accompanied by segmentation masks generated from three well-established handcrafted fire detection algorithms. Additionally, a secondary dataset with 9,044 manually annotated image patches provides human-verified references for model validation. The study evaluates deep learning models, particularly U-Net variants, to approximate and, in some cases, surpass the accuracy of traditional algorithms. It highlights the limitations of handcrafted algorithms, such as fixed-threshold constraints, while demonstrating the adaptability of deep learning-based fire detection. The research also explores ensemble methods, which outperform individual models. The dataset and models are publicly available, setting a benchmark for future fire detection studies, with potential applications in cross-satellite model transfer and spatiotemporal fire monitoring.

The FLOGA dataset is a machine-learning-ready resource for burned area mapping and severity estimation using Sentinel-2 imagery, with a focus on wildfire-prone regions in Greece [9]. Spanning 80 wildfire events (2017–2021), it integrates Copernicus Emergency Management Service (EMS) grading maps, providing manually annotated severity levels from 0 (unburned) to 4 (completely destroyed). The dataset supports binary burned area classification, severity assessment, and pre-/post-fire analysis, facilitating deep learning applications in wildfire detection and impact prediction. It has been preprocessed into 512×512 pixel patches at 10-meter resolution, ensuring high-quality inputs for training machine learning models. Benchmark evaluations using U-Net and SegFormer demonstrated the

effectiveness of multispectral bands and vegetation indices in improving severity classification. While challenges such as cloud cover and annotation subjectivity persist, future improvements aim to expand its geographic coverage, integrate Sentinel-1 data, and explore synthetic data augmentation. As an open-access dataset, FLOGA advances research in wildfire monitoring, severity mapping, and disaster response efforts.

The Wildfire-S1S2-Canada dataset [10] is a large-scale annotated resource for wildfire burned area mapping, integrating bi-temporal Sentinel-1 SAR and Sentinel-2 MSI data to support deep learning-based segmentation and change detection. It focuses on wildfire events in Canada (2017–2019), using data from the National Burned Area Database (Natural Resources Canada), with a 20-meter resolution for precise wildfire monitoring. The dataset includes pre-fire images, generated from a one-year time series before the fire season, and post-fire images, collected within two months of containment. Cloud-free Sentinel-2 imagery and rasterized wildfire perimeters were used to create high-quality reference masks, resulting in 9,052 training patches and 2,856 testing patches at 256×256 pixel resolution. It provides six Sentinel-2 spectral bands (RGB, NIR, SWIR-1, SWIR-2) and dual-polarization Sentinel-1 data (VH, VV) for multi-sensor analysis. Designed as a benchmark for deep learning-based wildfire mapping, this dataset supports advancements in near real-time fire progression monitoring and environmental impact assessment, with future expansions planned to increase spatial and temporal coverage for enhanced wildfire management applications.

Various U-Net models and their extensions have been widely applied in forest fire-related tasks. Adaptations such as attention mechanisms, residual blocks, and integration of temporal and environmental data have enhanced the effectiveness of U-Net in wildfire detection, smoke segmentation, and severity assessment.

Farasin et al. introduced a methodology for assessing wildfire damage severity using Sentinel-2 satellite imagery through a deep learning model named Double-Step U-Net [11]. In their study, they also proposed and compared the performance of Single, Parallel, and Double-Step U-Net frameworks, evaluating their effectiveness in burned area detection and severity estimation.

Monaco et al. utilized an extension of the Double-Step U-Net [11] to predict wildfire severity using a multichannel deep learning framework for post-fire satellite imagery analysis. Their approach, referred to as the Double-Step Deep Learning Framework (DSF) [12], incorporates an attention mechanism into the U-Net. The framework compares its results with derivatives of the U-Net model to evaluate

its effectiveness. The DSF operates in two steps: the first performs binary segmentation to identify burned regions, while the second assigns severity levels to each pixel within the identified burned areas. Leveraging Sentinel-2 data with twelve channels, the DSF captures diverse terrain features and applies multi-channel attention analysis to assess the importance of each spectral band in predicting burn severity.

MultiResUNet model, proposed by Ibtehaz and Rahman [13], is an enhancement of the original U-Net architecture designed specifically for improved performance in medical image segmentation. While traditional U-Net consists of an encoder-decoder structure with simple convolutional layers, MultiResUNet introduces MultiRes blocks. Furthermore, Ibtehaz and Rahman replaced the standard U-Net skip connections with Res paths. These paths apply additional convolutional layers to reduce the semantic gap between encoder and decoder features.

Han et al. proposed a transformer-based change detection model to accurately map burned areas and assess the severity of the burn [14]. Their approach leverages multispectral Landsat-8 imagery and auxiliary environmental data such as vegetation indices to enhance model performance and accuracy in identifying fire-affected areas. The model compares pre-fire and post-fire images to detect changes in vegetation and land surface features. This work demonstrates the potential of combining transformer-based methods with environmental data to improve wildfire impact assessments and aid in post-fire recovery planning.

Pinto et al. developed a deep learning model called BA-Net, which is tailored for identifying and timing burned regions by utilizing temporal series of satellite images from the VIIRS sensor [15]. This approach addresses key limitations of traditional burned area mapping techniques by eliminating the need for cloud masking and complex preprocessing steps. The model utilizes a U-Net architecture modified with a Long-Short-Term Memory (LSTM) layer to capture temporal dependencies.

Mengna Li et al. developed an early forest fire segmentation algorithm named F-Unet, designed to aid fire rescue operations [16]. F-Unet introduces a contraction path, a feature fusion layer, and an expansion path to enhance the precision of the segmentation. The contraction path leverages the first 13 layers of VGG16 to capture multiscale feature maps. The feature fusion network integrates these features to improve segmentation accuracy. The results from the experimental evaluation on the FLAME dataset indicate that F-Unet substantially improves the accuracy of fire segmentation.

Khennou and Akhloufi presented FU-NetCastV2, a deep learning convolutional neural network developed to predict fire spread and map burned areas [17]. FU-NetCastV2 identifies areas around wildfires that are at high risk of future spread by integrating satellite imagery with topographical and weather data. This approach leverages consecutive forest wildfire perimeters, satellite images, Digital Elevation Model maps, aspect, slope, and weather variables.

Wang et al. proposed an enhanced approach to early forest fire smoke detection through the development of Smoke-Unet, an improved U-Net model incorporating attention mechanisms and residual blocks [18]. Their method leverages the RGB, SWIR2, and AOD bands to improve recognition accuracy, demonstrating a 3.1% improvement over standard U-Net performance on a diverse dataset encompassing multiple seasons, regions, and types of land cover.

3
UNDERSTANDING WILDFIRES

3.1 Definition of Wildfires

Wildfires are uncontrolled fires that rapidly spread across vegetation-dense regions, including forests, grasslands, savannas, and scrublands. Unlike prescribed burns used for agricultural or land management purposes, wildfires ignite and propagate without direct human control, often posing significant threats to ecosystems and human settlements.

The scale and intensity of wildfires can vary considerably, ranging from localized brushfires to massive conflagrations that engulf thousands of hectares. Their spread depends on several factors, including fuel availability, weather conditions, wind speed, and topography.

Wildfires have become an increasingly critical environmental and societal challenge, with recent data indicating a rise in both frequency and severity. This escalation is driven by climate change, human-induced activities, and insufficient preventive measures, leading to devastating consequences for biodiversity, air quality, and public safety.

3.2 Causes of Wildfires

According to the General Directorate of Forestry (OGM), wildfires in Türkiye originate from four primary causes: negligence, intentional acts, natural causes, and unknown sources [19]. Fires caused by negligence typically result from human activities such as uncontrolled agricultural burning, discarded cigarettes, unattended campfires, and sparks from machinery. Intentional fires, including those started for land clearing, illegal activities, or disputes, contribute significantly to wildfire occurrences. The majority of wildfires are human-induced. Natural causes, particularly lightning strikes, also play a role in fire ignition. In some cases, the cause remains undetermined and is classified as unknown. Figure 3.1 presents a pie

chart illustrating the distribution of wildfire causes in Türkiye in the year 2023.

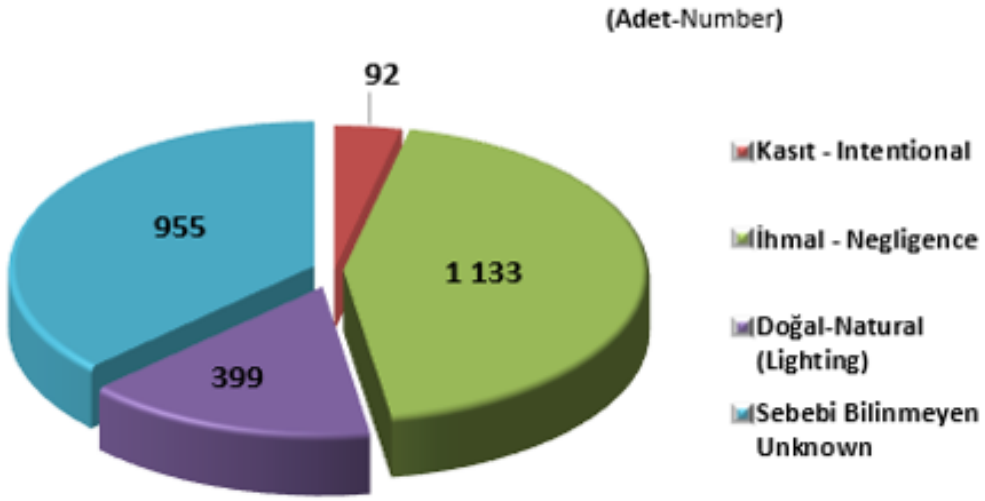


Figure 3.1 Total output according to cause of fire, 2023 [20]

The frequency of naturally occurring wildfires is expected to rise due to an increase in extreme weather events such as prolonged droughts, heatwaves, and stronger storm systems. These changes, largely attributed to climate variability and global warming, create conditions that facilitate fire ignition and spread, further intensifying wildfire risks.

3.3 Impacts of Wildfires

Wildfires have severe consequences on ecosystems, leading to significant biodiversity loss. The destruction of habitats forces wildlife to migrate, often into areas unsuitable for survival, disrupting ecological balance. Many species, particularly those with limited mobility, face population declines or even local extinction due to the rapid spread of fires. According to the European Commission's Joint Research Centre (JRC), wildfires in 2022 burned an area within the EU roughly equivalent to the size of Corsica, causing severe habitat destruction and threatening biodiversity [21].

Additionally, wildfires contribute to climate change by releasing vast amounts of greenhouse gases such as carbon dioxide (CO_2) and methane (CO_4) into the atmosphere. These emissions exacerbate global warming by increasing the concentration of heat-trapping gases, further intensifying extreme weather conditions [21]. Soil degradation is another critical consequence; intense fires strip the land of its protective vegetation cover, making it prone to erosion.

The economic toll of wildfires is substantial, affecting both individuals and governments. The destruction of infrastructure, including homes, power lines, and transportation networks, results in billions of euros in damages annually. In 2023, wildfires caused extensive economic damage across Southern Europe, requiring significant rebuilding efforts and disaster relief funding [21]. Fire suppression and post-fire recovery costs further strain national budgets, as governments allocate substantial resources to contain fires, rehabilitate burned areas, and provide financial aid to affected communities. The JRC's 2023 Wildfire Report emphasizes that economic losses extend beyond direct damages, impacting key industries such as agriculture, tourism, and forestry, which depend on healthy landscapes [21]. In regions prone to frequent wildfires, rising insurance claims drive up premiums, increasing the financial burden on homeowners and businesses [21].

Wildfires also pose a direct threat to human communities, often resulting in large-scale evacuations, displacement, and tragic loss of life. Entire neighborhoods can be reduced to ashes within hours, forcing families to relocate and rebuild from scratch. The emotional and psychological toll on survivors is immense, as they cope with the loss of homes, livelihoods, and loved ones. Beyond immediate displacement, wildfires have severe public health implications. According to the European Environment Agency (EEA), wildfire smoke contains fine particulate matter (PM2.5) and toxic gases that exacerbate respiratory illnesses, particularly among vulnerable populations such as children, the elderly, and individuals with preexisting conditions [22]. Fine particulate matter is one of the most harmful air pollutants to human health and a significant component of wildfire smoke. Long-term exposure to wildfire smoke has been linked to chronic respiratory diseases, cardiovascular issues, and increased mortality rates. The growing intensity and frequency of wildfires highlight the urgent need for improved mitigation strategies, stricter land management policies, and enhanced community preparedness efforts across Europe [22].

3.4 Wildfire Trends

The wildfire statistics in Türkiye from 2020 to 2023 reveal significant variations in both the number of fires and the burnt area. The year 2021 stands out as the most destructive, with 139,503 hectares burned, which is over ten times the burnt area in 2022 and 2023. Despite having fewer total fires (2,793) compared to 2020 (3,999), the fires in 2021 caused much greater devastation, highlighting the intensity of the incidents.

According to the General Directorate of Forestry, negligence and accidents were the leading contributors across all years, with 2021 having the highest burnt area (46,879 ha) due to such incidents. Fires of unknown origin also accounted for a significant portion of burnt land, particularly in 2021 (46,269 ha). Intentional fires, though fewer in number, resulted in substantial damage in 2021 (46,147 ha) but declined sharply in subsequent years. Lightning-induced wildfires were the least frequent, with a moderate increase in 2023 (399 cases, 365 ha burned) compared to 2020 (312 cases, 197 ha burned) [20].

Table 3.1 Wildfire Statistics in Türkiye (2020-2023) [20]

Year	Burnt Area (ha)	Total Fires	Intentional		Negligence-Accident		Natural (Lightning)		Unknown	
			Fires	Area (ha)	Fires	Area (ha)	Fires	Area (ha)	Fires	Area (ha)
2020	20,971	3,999	72	718	1,156	8,285	312	197	1,859	11,771
2021	139,503	2,793	110	46,147	1,001	46,879	353	208	1,329	46,269
2022	12,799	2,160	86	4,722	830	5,428	358	517	886	2,132
2023	15,520	2,579	92	663	1,133	11,259	399	365	955	3,233

In 2024, the United States experienced a significant wildfire season. According to the National Interagency Fire Center (NIFC) [23], there were 61,685 wildfires, which burned approximately 8,851,142 acres nationwide. While the number of fires was the seventh lowest since 2000, the total acreage burned was the seventh highest, indicating larger and more intense fires on average. The 2024 wildfire season in the United States, particularly in California, highlights the escalating challenges posed by wildfires. The combination of climate change, increased fuel loads, and expanding human settlements in fire-prone areas necessitates a comprehensive approach to wildfire management, such as prevention, early detection, and effective response strategies.

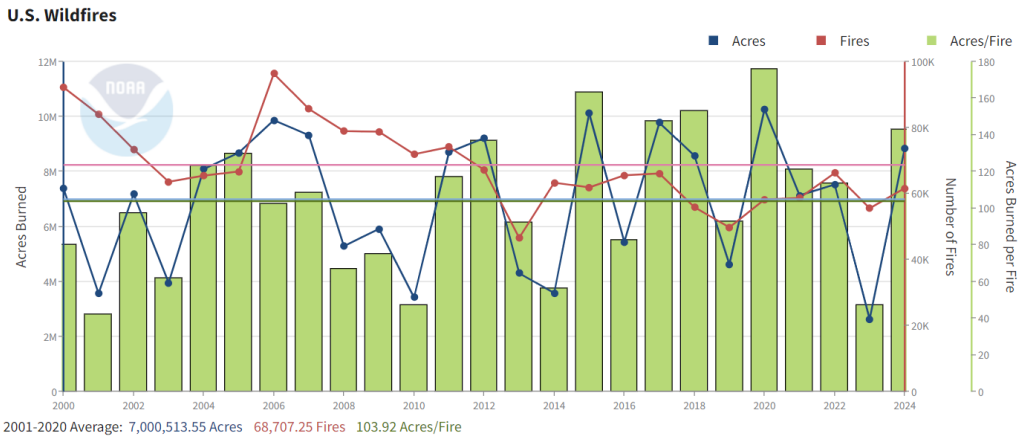


Figure 3.2 US wildfires since 2020 [23]

3.5 Post-Fire Land Management Strategies and Severity Assessment

Effective post-wildfire management strategies vary across regions but generally focus on stabilizing landscapes, preventing secondary disasters such as soil erosion and flooding, and restoring biodiversity. Different countries have developed specialized frameworks to address post-fire recovery based on their unique ecological and climatic conditions. Among the various approaches, Türkiye, the United States, and Europe have implemented region-specific recovery programs tailored to their environmental and policy needs.

In Türkiye, post-wildfire recovery and rehabilitation efforts are primarily managed by the General Directorate of Forestry (OGM). Following a wildfire, OGM promptly initiates reforestation and rehabilitation projects to restore the affected areas. These initiatives include natural regeneration, afforestation, erosion and flood control, and enhancement of natural vegetation. For instance, in the Yamanlar region, a comprehensive program was implemented to reforest and rehabilitate burned forest areas, utilizing species such as holm oak, cork oak, hairy oak, stone pine, and black pine [24].

In the United States, the Burned Area Emergency Response (BAER) [25] program is a key initiative that addresses post-fire emergency stabilization to protect public safety and prevent further degradation of the landscape. BAER teams assess immediate risks such as soil erosion, flooding, and threats to human life and property, implementing treatments like mulching, seeding, and installing erosion barriers to mitigate these hazards. A critical component of this assessment is the creation of a Soil Burn Severity map, which is developed using satellite imagery and validated through field surveys. This map categorizes areas into low/unburned, moderate, and high burn severity, providing essential information for understanding watershed responses to wildfires. In post-wildfire scenarios, agencies like the BAER teams in the United States utilize Soil Burn Severity (SBS) maps to inform their stabilization and rehabilitation strategies. These maps categorize affected areas into unburned/low, moderate, and high burn severity classes, guiding the selection of appropriate treatments. Treatment strategies based on burn severity are listed below.

1. Unburned/Low Severity Areas:
    - Natural Recovery: These areas often retain sufficient vegetation and organic matter, allowing ecosystems to recover naturally without intervention.

- Monitoring: Regular assessments ensure that recovery is progressing as expected and that no unforeseen issues arise.

2. Moderate Severity Areas:

    - Erosion Control Measures: Implementing treatments such as mulching or seeding with native species helps stabilize the soil and prevent erosion.
    - Invasive Species Management: Active monitoring and control of invasive species are crucial to support the reestablishment of native vegetation.

3. High Severity Areas:

    - Immediate Stabilization: Applying mulch, installing erosion barriers, or constructing check dams can prevent significant soil loss and manage water runoff.
    - Reforestation Efforts: Active planting of native tree and shrub species accelerates habitat restoration and reduces the risk of invasive species colonization.

These targeted approaches, informed by burn severity assessments, are essential for effective post-wildfire recovery, ensuring that interventions are appropriately matched to the specific conditions of each affected area.

Furthermore, the Burn Severity Portal, an interagency platform, provides comprehensive access to federal burn severity mapping data and field assessments. This portal offers information about various burn severity mapping programs and access to current and historical data products, supporting informed decision-making in post-fire recovery efforts.

Beyond the United States, European countries employ different strategies focused on mitigating post-wildfire soil erosion and promoting ecosystem resilience. Post-wildfire recovery and rehabilitation efforts are multifaceted, focusing on mitigating immediate environmental impacts and promoting long-term ecosystem resilience. A significant concern following wildfires is soil erosion, which can increase sharply, leading to further environmental degradation. Studies have shown that post-fire mitigation measures can reduce soil erosion impacts by 63–77%, highlighting the importance of timely interventions [26].

While Türkiye emphasizes afforestation, the U.S. focuses on immediate stabilization, and Europe prioritizes erosion control and long-term ecosystem

resilience. These diverse approaches reflect the need for adaptive post-wildfire recovery strategies based on local environmental conditions and risk factors.

4
MATERIAL AND METHOD

In this chapter, the methodologies that were proposed and worked on in this thesis will be reviewed and discussed, In addition, the materials including data and programs used in this work.

4.1 Dataset

The dataset structure and fold separation follow the methodology outlined in the study "*A dataset for burned area delineation and severity estimation from satellite imagery*" [1]. It comprises data from 73 distinct forest wildfire events recorded between 2017 and 2019 across various European regions, with a total monitored area of approximately 19,000 km². Each wildfire event is represented by four key images: pre-fire and post-fire Sentinel-1 and Sentinel-2 imagery, along with a mask image denoting the burned area. The mask images are encoded in grayscale to indicate five levels of severity, ranging from level 0 (unburned) to level 4 (completely destroyed).

The dataset was constructed using data from the Copernicus Emergency Management Service (EMS) Rapid Mapping products and the Sentinel Hub platform. The temporal window for data collection was standardized to include images from one month prior and one month following the wildfire event. This ensures the availability of sufficient pre- and post-fire data for analysis. The dataset also includes a validity mask to indicate whether pixel data in the satellite images are valid or cloud-covered.

Table 4.1 The organization schema of the validation and test sets in the experiments

Sets	**Folds**						
Test	Purple	Coral	Pink	Grey	Cyan	Lime	Magenta
Validation	Coral	Cyan	Coral	Coral	Coral	Coral	Coral

The fold colors represent sets grouped by their similar morphological features, primarily corresponding to regions in closely situated cities throughout Europe. As shown in Table 4.1, if the test set is Purple, the validation set will be Coral. The remaining folds form the training set. The majority of the dataset consists of unburned areas, as illustrated by the logarithmic representation in Figure 4.1. Due to the overwhelmingly high number of unburned pixels, a logarithmic scale was used to enable better visualization and interpretation of the data distribution. This figure also reveals that the Pink fold dataset contains very few pixels at level 4, indicating severely burned regions are scarce in that fold. In total, the dataset includes 73 images, with the number of images per fold as follows: Purple (8), Coral (12), Pink (14), Grey (9), Cyan (11), Lime (10), and Magenta (9). This disparity highlights the difficulty of training models effectively to predict burned areas.

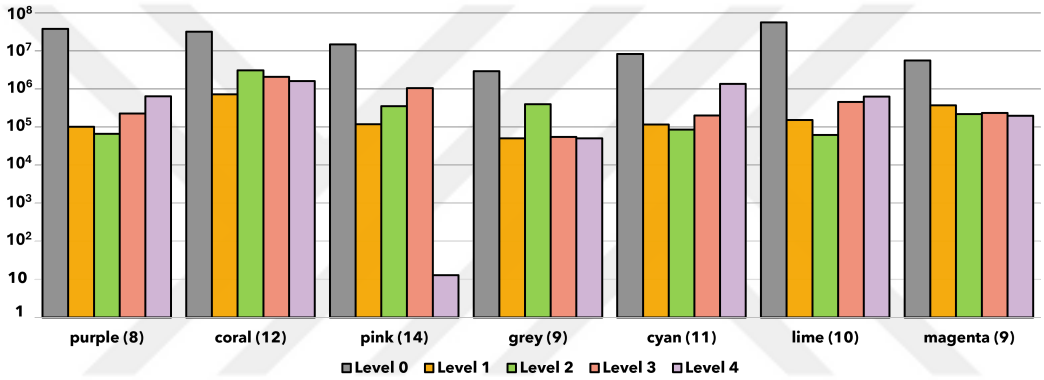


Figure 4.1 Logarithmic representation of pixels from mask images based on folds

This research exclusively utilized post-fire Sentinel-2 imagery for the experiments. During the selection process, images with excessive cloud cover were filtered out by applying a cloud cover threshold. The resolution of the images is excessively high for straightforward use in training neural network models. Consequently, the images were split into square tiles, each measuring 480x480 pixels.

During training, random data augmentation techniques were applied to improve model performance, ensure robustness, address data imbalances, and prevent overfitting. The methods employed include rotating images by a maximum of 50 degrees, cropping images by as much as 20 degrees, and flipping images both horizontally and vertically. Each of these techniques was applied with a 50% probability for each batch of images. Furthermore, all images were normalized using a mean and standard deviation of 0.5. Only the training dataset is enhanced using data augmentation methods like rotation and flipping.

4.2 Double-Step Deep Learning Framework [12]

Double-Step Deep Learning Framework is designed to predict the severity of wildfires, dividing the task into two steps: wildfire detection and severity estimation. Figure 4.2 outlines the general structure of the Double-Step Framework (DSF).

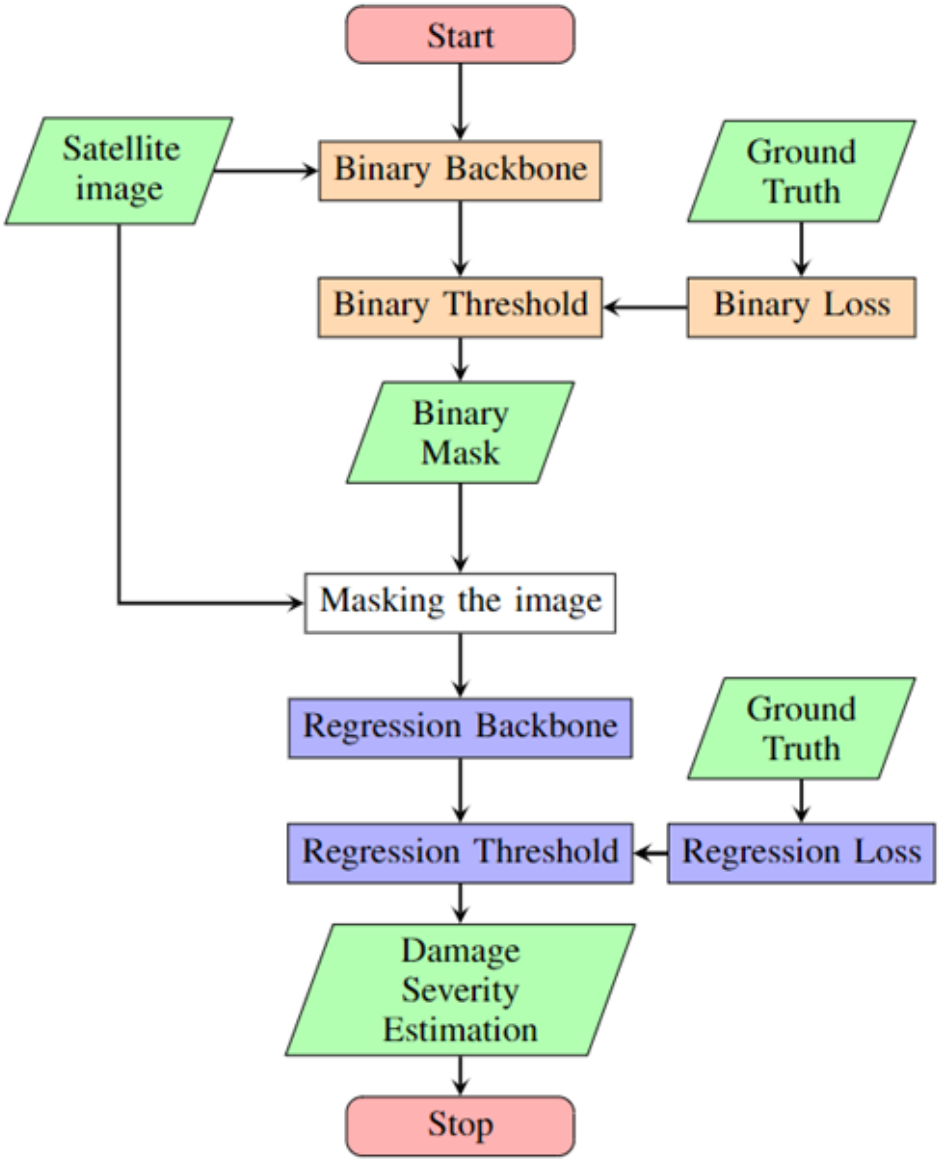


Figure 4.2 Double-Step framework architecture [12]

4.2.1 Step 1: Wildfire Detection

The initial phase of the DSF, known as the Wildfire Detection Task (Step One), aims to identify if a particular region has been burned. In this stage, a binary-class backbone was utilized that analyzes satellite images and generates a binary probability map. Each pixel represents the likelihood of being burned (values between 0 and 1). The backbone is trained using a binary loss function to

improve segmentation accuracy. This binary segmentation step is trained using Binary Cross-Entropy with Logits Loss (BCEWithLogitsLoss), which combines the sigmoid activation function and binary cross-entropy loss into a single operation. This formulation prevents numerical instability and improves training efficiency.

After generating the probability map, the values are thresholded to create a final binary mask (burned/unburned regions). This research utilized a threshold value of 0.5. Applying the resulting binary mask to the original image effectively conceals the unburned regions, ensuring that only the burned areas are visible in preparation for the model's subsequent step.

4.2.2 Step 2: Severity Estimation

In the second step, the model uses the masked satellite images and predicts the severity of fire damage using a regression backbone. The result is a map assigning each pixel a severity score from 0 to 4, indicating varying levels of burn damage. This step is formulated as a regression problem. The Mean Squared Error (MSE) loss function is employed to compute the average squared difference between predicted severity values and ground truth. If the binary mask produced initially contains inaccuracies, it will negatively impact the performance of the second model. Thus, ensuring an accurate binary mask enhances the overall effectiveness of severity estimation.

The second step in the Double-Step Deep Learning Framework (DSF) is considered a regression problem because it involves estimating continuous severity levels for burned areas, ranging from 0 (no damage) to 4 (completely destroyed). The regression backbone does not use an activation function in its output layer, allowing the model to produce continuous values within the range [0, 4]. The output is then clipped and rounded to the nearest integer to match the discrete severity levels. The regression model operates only on burned areas identified by the binary detection backbone in the first step. The binary mask filters unburned areas, reducing data variability and enabling the regression model to focus on accurately estimating severity levels for burned regions.

The Double-Step Framework approach supports the implementation of various neural network architectures for both binary detection and severity assessment processes. In this study, several deep learning-based segmentation networks have been deployed. Each step uses the same deep learning model architecture. The subsequent subtitles provide a summary of the models employed.

4.2.3 Implemented Models

The Double-Step Framework incorporates four U-Net-based models (U-Net, Attention U-Net, Nested U-Net, and SegU-Net) to address burned area detection and severity assessment. Each model leverages unique architectural adaptations to enhance segmentation performance and robustness in complex wildfire scenarios.

The following sections provide detailed explanations of the design and functionality of these models.

4.2.3.1 U-Net

U-Net [27] is a convolutional neural network architecture specifically designed for image segmentation tasks. As seen in Figure 4.3, the downsampling and upsampling operations form a U-shaped structure, which characterizes the U-Net architecture. The architecture consists of two main components: an encoder and a decoder.

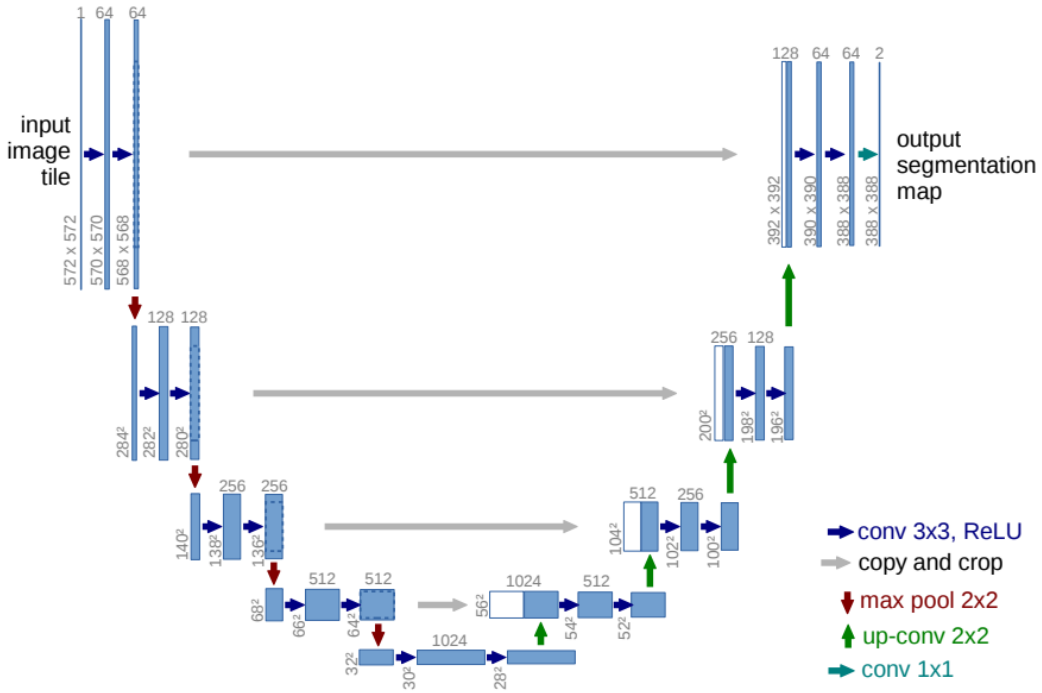


Figure 4.3 U-Net architecture [27]

The encoder, also known as the contracting path, captures the context and essential features of the input image. It consists of a sequence of convolutional layers followed by pooling operations. These convolutional layers use small receptive fields (e.g., 3×3 filters) to detect local patterns while preserving spatial resolution. The pooling operations progressively downsample the feature maps, reducing their spatial dimensions but increasing the depth of the feature representations.

This hierarchical representation allows the model to learn complex patterns and contextual information at different scales.

The decoder, or expansive path, is tasked with reconstructing spatial details and creating high-resolution feature maps from the compressed representations learned by the encoder. This process involves upsampling operations, typically performed using transposed convolutions or bilinear interpolation, followed by convolutional layers. The decoder progressively restores the spatial resolution of the feature maps, resulting in a final output that matches the size of the input image.

One of U-Net's distinguishing features is the incorporation of skip connections between corresponding layers in the encoder and decoder. These skip connections allow the model to combine low-level spatial features from the encoder with high-level semantic information from the decoder. By directly linking features at the same resolution level, skip connections help preserve fine details and mitigate the loss of spatial information that can occur during downsampling.

U-Net has been extensively used for tasks such as tumor segmentation, organ delineation, and cell tracking. Its ability to handle limited data and produce highly accurate results makes it ideal for medical datasets. In remote sensing, U-Net has proven effective for segmenting features like burned areas, water bodies, and vegetation from multi-channel satellite imagery.

Since its introduction, U-Net has inspired numerous variations and extensions, such as Attention U-Net and MultiResUNet, which enhance its performance by integrating attention mechanisms or multi-resolution features. Its success demonstrates the power of convolutional architectures for segmentation tasks and highlights the importance of designing models tailored to specific challenges.

Overall, U-Net remains a foundational architecture in the field of image segmentation, valued for its simplicity, versatility, and effectiveness in addressing diverse challenges across domains.

4.2.3.2 Attention U-Net

Attention U-Net [28] is an advanced version of the original U-Net architecture, designed to further improve segmentation accuracy by incorporating attention mechanisms. This extension builds upon U-Net's strengths by integrating attention gates into the skip connections. These gates enable the model to selectively focus on relevant regions of the image while suppressing less important or noisy areas, ensuring that only the most critical features are passed to the decoding layers.

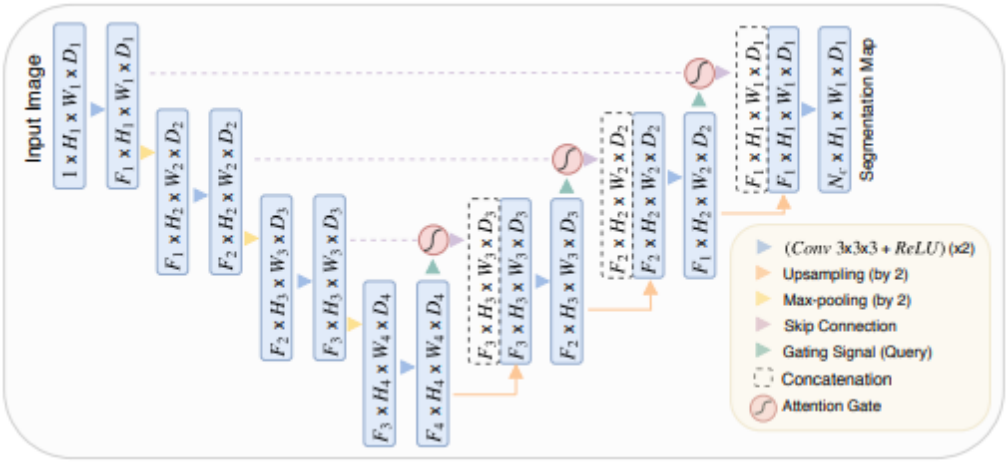


Figure 4.4 Attention U-Net architecture[28]

Attention gates operate by analyzing feature maps from both the encoder and the decoder paths. They assign weights to different regions of the input based on their importance. This highlights areas that are more likely to contribute to accurate segmentation. This selective emphasis allows the model to prioritize significant features while discarding irrelevant information. The attention mechanism is particularly beneficial for tasks involving images with high variability or complex backgrounds.

Despite the addition of attention gates, Attention U-Net retains the same overall encoder-decoder structure as the original U-Net. The encoder captures contextual information, while the decoder reconstructs the spatial details. However, the inclusion of attention gates refines the process by ensuring that the decoder receives more precise and contextually relevant features. In Figure 4.4, the architecture of Attention U-Net is shown.

Figure 4.5 illustrates the structure of the attention gate. The attention gate processes feature maps from both the encoder (X^l) and the decoder (g) paths. Here, X^l represents the low-level spatial features from the encoder, while g provides high-level contextual information from the decoder. The attention mechanism aligns these features through transformations using linear weights, W_x and W_g. The output of these transformations is combined via element-wise addition to produce an intermediate feature map. This map is activated by the ReLU function (σ_1) and passed through another convolution (ψ). The sigmoid function (σ_2) normalizes the attention coefficients (α) to values between 0 and 1. These coefficients are then used to resample the original feature map X^l by performing an element-wise multiplication. The refined feature map ($\hat{X}^l$) is subsequently forwarded to the decoder via the skip connection. This process enables the model to focus on the

most relevant regions for segmentation.

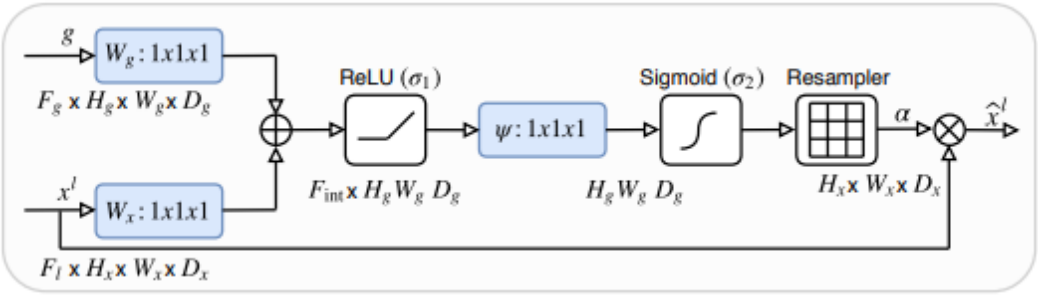


Figure 4.5 Attention gate[28]

Attention U-Net offers several key benefits that enhance its performance in segmentation tasks. It improves localization by enabling the model to detect small or intricate structures within images more effectively. The attention mechanism also helps suppress noise by filtering out irrelevant features, resulting in cleaner segmentation masks. Additionally, attention gates dynamically adjust their focus based on the input image, allowing the model to adapt to different scenarios. Despite these modifications, Attention U-Net retains the flexibility and adaptability of the original U-Net, making it suitable for a wide range of applications.

4.2.3.3 Nested U-Net

Nested U-Net[29], also known as UNet++, is an enhancement of the original U-Net architecture designed to improve segmentation performance through the introduction of multiple nested and dense skip pathways between the encoder and decoder. This architecture introduces a series of intermediate layers called nested dense skip connections, which refine feature propagation and improve the network's ability to capture both low-level and high-level features.

In Nested U-Net, the traditional skip connections in U-Net are replaced by a dense series of nested connections that connect encoder and decoder blocks at varying depths. These connections enable a more granular fusion of feature maps from different levels of the network, enhancing the model's ability to learn hierarchical representations. By passing features through these intermediate layers, the model reduces the semantic gap between the encoder and decoder paths, resulting in better segmentation outputs.

The Figure 4.6 illustrates the Nested U-Net architecture. The encoder backbone extracts hierarchical features at different resolutions ($X^{i,0}$), while the decoder path progressively reconstructs the segmentation map ($X^{i,j}$) through upsampling and convolutional refinement. Dense skip connections (dotted arrows) transfer

information between the encoder and decoder levels. This mechanism enables the integration of fine-grained spatial details and high-level contextual features. Each circle represents a convolutional operation. Supervision is applied at multiple refinement levels ($\mathcal{L}$) to encourage accurate intermediate predictions. The final segmentation map ($X^{0,4}$) combines refined features.

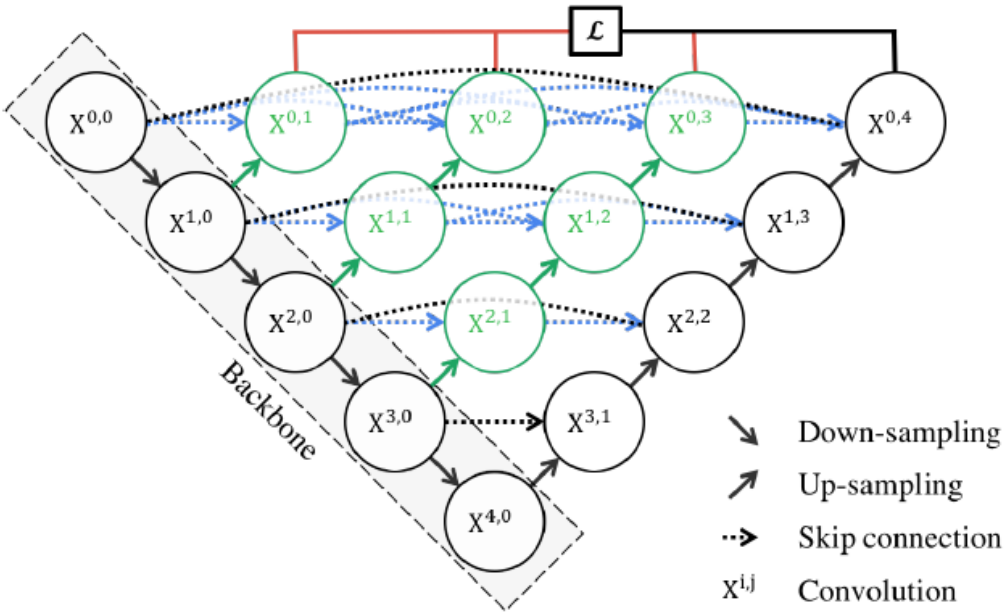


Figure 4.6 Nested U-Net architecture[29]

The Nested U-Net architecture offers several benefits that enhance its performance in segmentation tasks. The dense skip pathways enable improved feature fusion, allowing for better integration of features across different scales and leading to more precise segmentation. Its nested structure facilitates the capture of complex patterns and hierarchical relationships within the input data. Additionally, Nested U-Net is highly flexible and can be adapted to various applications, including medical imaging and satellite image analysis. The incorporation of additional layers and dense connections contributes to significant performance gains. These improvements are especially evident in tasks involving complex structures or fine details.

4.2.3.4 SegU-Net

SegUNet [30] is a U-Net variant that combines the original U-Net architecture with a pretrained VGG backbone for feature extraction. This approach leverages the robust feature extraction capabilities of VGG, a well-known convolutional neural network designed for image classification, to enhance segmentation performance in tasks requiring detailed and accurate masks.

SegUNet features an encoder-decoder structure where the encoder uses a pretrained VGG network to process the input image and extract detailed spatial and contextual features. It retains U-Net's characteristic skip connections. These connections combine low-level spatial features from the encoder with high-level semantic information in the decoder. The skip connections enhance the model's ability to generate detailed and precise segmentation masks, particularly for tasks involving complex structures or fine boundaries.

The pretrained VGG encoder provides a strong initialization, particularly beneficial when working with small datasets. It transfers learned features from large-scale datasets to the segmentation task. This pretraining improves both convergence speed and accuracy.

SegUNet offers several benefits that enhance its performance in segmentation tasks. The use of a pretrained VGG encoder improves performance and generalization even when training data is limited. By retaining skip connections, the model preserves fine spatial details and ensures accurate reconstruction in the decoder. SegUNet is highly adaptable for various segmentation tasks, including medical imaging and remote sensing.

4.3 Proposed Methods

This thesis builds upon the framework proposed by Monaco et al. [12], which utilizes a double-step approach for wildfire detection and severity assessment. The contributions of this work lie in the integration of the MultiResUNet model, experimentation with U-Net variants, and evaluation of the framework under different configurations. Additionally, Mask R-CNN is employed as a baseline model for binary segmentation using RGB channels. The experiments were conducted on pre-processed satellite imagery, retaining all 12 spectral channels for UNet-based models.

The study further explored the effect of severity mask configurations by testing both Level 4 and Level 5 masks. Reducing the severity masks from five to four levels aimed to evaluate the trade-off between computational efficiency and the granularity of severity predictions. Unlike the U-Net variants, Mask R-CNN was trained using only the RGB channels of the satellite images with pre-trained weights. This setup demonstrated the limitations of using three-channel imagery in wildfire detection tasks. It emphasized the importance of incorporating multi-spectral data in advanced models like MultiResUNet. The subsequent sections detail the models and experiments.

4.3.1 Mask R-CNN

Mask R-CNN [31] is an advanced deep learning architecture designed for instance segmentation tasks. It extends the capabilities of Faster R-CNN [32] by incorporating an additional branch for predicting segmentation masks on each detected object. This makes Mask R-CNN a hybrid model capable of performing both object detection and pixel-level segmentation, making it highly effective for tasks requiring precise localization.

Mask R-CNN begins with a backbone network, such as ResNet [33], for feature extraction. The extracted features are passed to a Region Proposal Network (RPN), which generates candidate regions likely to contain objects. These candidate regions are refined using RoI Align, a crucial operation that ensures precise alignment of proposed regions to the feature map. This refinement is essential for accurate bounding box and mask predictions. For each region, Mask R-CNN predicts the object class, bounding box, and a binary mask that delineates the object at the pixel level. In Figure 4.7, the architecture of Mask R-CNN is shown.

The mask prediction branch operates independently of the class and bounding box predictions, ensuring high flexibility and accuracy. This multi-task learning approach allows Mask R-CNN to excel in segmenting objects in complex images with diverse object scales and shapes.

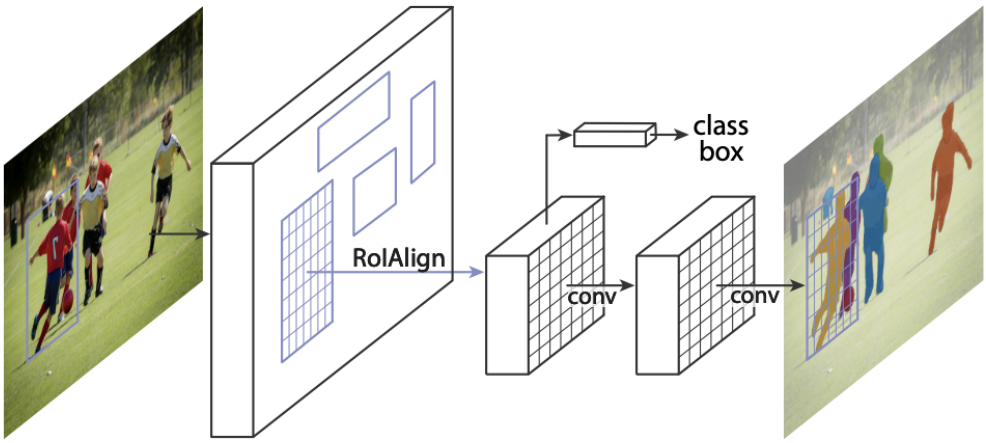


Figure 4.7 Mask R-CNN architecture[31]

Mask R-CNN offers several benefits that make it a powerful model for segmentation tasks. It provides comprehensive functionality by simultaneously performing object detection, classification, and instance segmentation. This allows the model to handle multiple vision tasks within a single framework. The model is highly flexible and can handle objects of various shapes, scales, and occlusions. It

achieves this with high precision, ensuring accurate segmentation and detection. Its backbone versatility allows it to support different architectures, such as ResNet and ResNeXt. Additionally, its high precision in pixel-level mask predictions enhances segmentation accuracy, making it particularly suitable for applications that require fine detail.

Mask R-CNN has diverse applications across various domains. In medical imaging, it is used for the precise localization and segmentation of tumors, organs, and anomalies in medical scans. In satellite image analysis, it effectively detects and segments features such as buildings, water bodies, and vegetation in high-resolution imagery. The model also plays a crucial role in autonomous driving, where it identifies and segments vehicles, pedestrians, and road signs in real-time scenarios. Additionally, in industrial inspection, it is employed for quality control tasks, enabling the detection and segmentation of defects in manufactured products.

4.3.2 MultiResUNet

MultiResUNet [13] is an advanced version of the U-Net architecture specifically designed for medical image segmentation. Unlike the classical U-Net, MultiResUNet incorporates MultiRes blocks, which are designed to capture image features at multiple resolutions simultaneously. This ability to process features at various scales allows the model to more effectively identify structures of different sizes within medical images, enhancing segmentation accuracy. The MultiResUNet architecture is shown in Figure 4.8, illustrating its similarity to the U-Net architecture with the addition of MultiRes blocks and Res Paths."

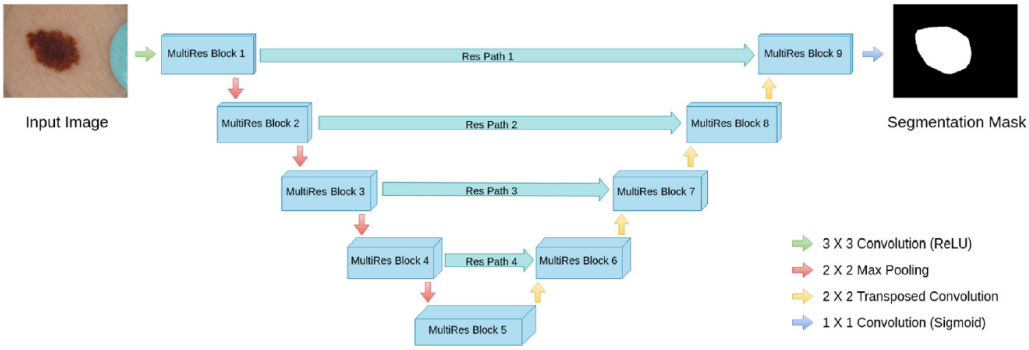


Figure 4.8 MultiResUNet architecture[13]

- MultiRes Blocks: These blocks integrate multiple convolutional layers with varying kernel sizes, allowing the network to extract features at different spatial resolutions. Each block performs parallel 3×3 convolutions to detect fine-grained details, while stacked filters capture deeper, high-level

representations. Additionally, a 1×1 convolution is applied to the input to ensure dimensional compatibility for the residual connection. The outputs of the multi-scale convolutions are concatenated to create a richer feature representation. These concatenated features are then combined with the residual connection, preserving the original spatial information while enhancing the overall feature set. This design is particularly beneficial in medical imaging tasks, where detecting both subtle and coarse structural patterns is crucial for accurate segmentation. In Figure 4.9, the architecture of the MultiRes block is shown.

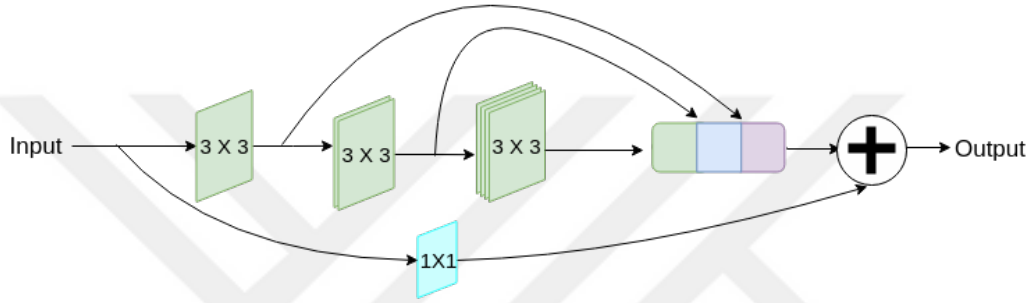


Figure 4.9 MultiRes block architecture[13]

- MultiResUNet replaces the standard skip connections with Res Paths, which are composed of multiple 1×1 convolutional layers followed by residual connections. These Res Paths help bridge the semantic gap between the encoder and decoder feature representations by progressively refining the transferred features. Each Res Path ensures that the feature maps are properly aligned in terms of depth and context before being combined in the decoder. As shown in Figure 4.10, the residual connections in the Res Paths preserve original information while allowing for incremental adjustments at each convolutional layer. This design improves the flow of information across the network and reduces the risk of degradation during reconstruction, particularly in deeper architectures. By refining the features step by step, Res Paths enhance the model's ability to reconstruct detailed and accurate segmentation outputs.

Although initially developed for medical imaging, MultiResUNet has been considered for tasks such as wildfire segmentation in satellite imagery. Wildfires present complex spatial patterns that vary significantly across geographic regions. MultiResUNet's ability to process multi-resolution features aligns well with the

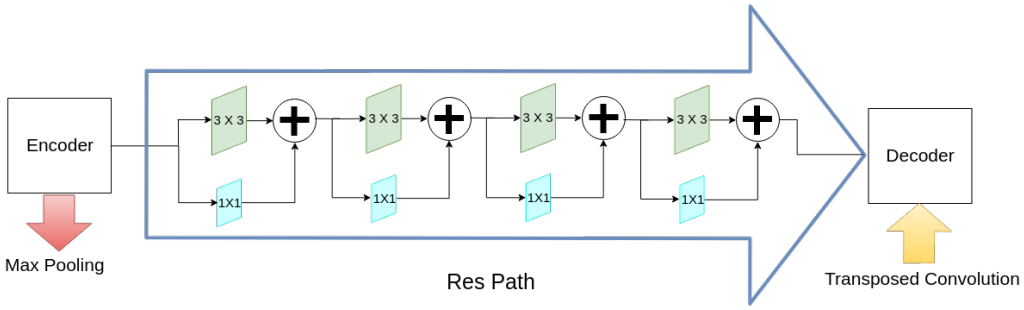


Figure 4.10 Res path architecture[13]

need for precise segmentation of burned areas. Its unique Res paths also make it particularly effective in bridging disparities between encoder and decoder features, enabling accurate segmentation of challenging environmental features.

MultiResUNet offers several benefits that make it a powerful architecture for segmentation tasks. The MultiRes blocks enhance the model's ability to extract multi-scale features. This enables the model to analyze and segment structures of varying sizes within an image. The inclusion of Res Paths improves information flow by reducing the loss of fine details during reconstruction, resulting in more accurate segmentation outputs. Its versatile design makes MultiResUNet highly suitable for a wide range of applications, such as medical imaging and satellite-based environmental monitoring. Furthermore, its robustness to variations in spatial patterns ensures adaptability across diverse segmentation tasks.

4.3.2.1 Implementing MultiResUNet to DSF

To integrate MultiResUNet into the Double-Step Framework, certain architectural adjustments were made to accommodate its role in the two-step process of wildfire segmentation and severity assessment.

- In the first step of the DSF, where the task is binary segmentation (burned vs. unburned areas), the last convolutional layer of MultiResUNet includes a sigmoid activation function. This enables the model to output a probability map indicating the likelihood of each pixel belonging to a burned area.
- In the second step, where severity levels are predicted (a regression task), the sigmoid activation is bypassed to allow the output layer to produce continuous values. This ensures the model can predict severity levels without constraints imposed by the activation function.

4.3.2.2 MultiResUNet Experiment Using Combined Pre- and Post-Fire Imagery

To evaluate the potential benefits of incorporating temporal information, an additional experiment was conducted using both pre-fire and post-fire satellite imagery. In this configuration, the pre-fire and post-fire images were concatenated along the channel dimension, resulting in a 24-channel input (12 channels for each timestamp). The same MultiResUNet architecture was used without modifications to the network structure, ensuring a fair comparison between configurations. This approach aimed to investigate whether including spectral differences from before and after the fire event could improve the model's ability to detect burned areas. The intuition behind this experiment lies in leveraging the temporal contrast to highlight fire-induced changes that may not be as distinct in post-fire imagery alone. The framework utilized in this study includes native support for using both pre-fire and post-fire images.

4.3.2.3 MultiResUNet Experiment with Decoder Batch Normalization Removed

To assess the impact of decoder normalization on performance, an alternative version of the MultiResUNet was implemented by removing the final batch normalization layer from each decoder block. This modification was designed to evaluate whether normalization in the later stages of the network was beneficial or potentially suppressing important fine-grained details necessary for precise segmentation. Batch normalization is a technique commonly used in deep learning to stabilize and accelerate training by normalizing the inputs of each layer. While it often improves convergence and generalization, it can also inadvertently smooth out spatial details that are critical for dense prediction tasks like segmentation. The rest of the network, including the encoder path and ResPath connections, remained unchanged. This lightweight architectural change maintains the core structure of the original MultiResUNet while offering a focused study.

4.3.3 Level 4 and Level 5 Experiments

The severity assessment component of this thesis explores the impact of using Level 4 and Level 5 severity masks on model performance. These experiments aim to evaluate the trade-off between the granularity of severity classification and the computational efficiency of the models within the framework.

The motivation behind experimenting with both Level 4 and Level 5 severity masks stems from the need to balance the complexity of severity classification

with practical considerations of model performance and computational efficiency. Wildfire damage varies across geographic regions and vegetation types, leading to a spectrum of severity levels that can be challenging to classify accurately.

- Level 5 Masks: These masks divide the severity of burned areas into five distinct levels, offering a finer granularity for severity assessment. The levels correspond to varying degrees of fire damage, ranging from unburned regions to severely burned areas. While providing detailed insights, this configuration increases the complexity of the regression task, as the model must learn to distinguish subtle differences between adjacent levels.
- Level 4 Masks: By consolidating the severity levels into four categories, this configuration reduces the granularity, simplifying the regression task. The reduction aims to enhance computational efficiency and improve model convergence, particularly for smaller datasets or limited computational resources. However, it sacrifices some detail in severity classification, which may impact the ability to differentiate between high-severity regions.

4.3.4 Hyperparameters

The following hyperparameters were used in training the DSF:

- **Batch size:** The number of training samples processed together in one forward and backward pass. A batch size of 8 was chosen to balance memory efficiency and gradient estimation stability.
- **Input channels:** The number of channels in the input images. Since the dataset consists of Sentinel-2 satellite images with 12 spectral bands, the input channels are set to 12 to fully utilize all available spectral information.
- **Epochs:**The total number of times the entire training dataset is passed through the model. A maximum of 50 epochs was selected to ensure sufficient learning while preventing excessive overfitting.
- **Seed:** A fixed random seed value (1) ensures reproducibility of results by maintaining consistency in dataset shuffling, weight initialization, and augmentation.
- **Patience:** The number of consecutive epochs without improvement before early stopping is triggered. Patience = 5 means training will halt if validation performance does not improve for 5 consecutive epochs.

- **Tolerance:** The minimum required improvement in validation loss to reset the early stopping counter. A tolerance of 0.01 prevents the model from stopping prematurely due to minor fluctuations.
- **Criterion:** The loss function used for binary segmentation. nn.BCEWithLogitsLoss is applied, as it combines binary cross-entropy loss with a sigmoid activation function, making it suitable for burned area classification.
- **Regr_criterion:** The loss function used for severity regression. Mean Squared Error measures the average squared difference between predicted and actual severity levels, making it suitable for continuous severity estimation.
- **Optimizer:** Adam optimizer is used for weight updates, with a learning rate of 0.0001 to ensure smooth and stable convergence. Adam is widely used due to its adaptive learning rate and momentum properties.
- **Trainer - Precision:** The numerical precision used in training. Precision = 16 enables mixed-precision training, which reduces memory consumption and speeds up training on GPUs.
- **Trainer - Gradient_clip_val:** The maximum allowable gradient norm (5.0) during backpropagation. This prevents gradient explosion, ensuring stable training, especially when using Adam optimizer.

Table 4.2 Hyperparameter settings

Hyperparameter	**Value**
Batch size	8
Input channels	12
Epochs	50
Seed	1
Patience	5
Tolerance	0.01
Criterion	`nn.BCEWithLogitsLoss`
Regr_criterion	`nn.MSELoss`
Optimizer	`optim.Adam (lr=0.0001)`
Trainer - Precision	16
Trainer - Gradient_clip_val	5.0

In Table 4.2, the hyperparameter settings are represented. A batch size of 8 was chosen to balance memory efficiency and stable gradient updates. The input data consisted of 12 spectral channels from Sentinel-2 post-fire imagery. The training

spanned 50 epochs, with an early stopping mechanism applied using a patience value of 5. This means that training would halt if no improvement was observed in validation loss for five consecutive epochs. Additionally, a tolerance threshold of 0.01 was set, requiring validation loss to improve by at least this amount for training to continue.

The model was trained with a learning rate of 0.001. The Adam optimizer was employed. Two distinct loss functions were used for the Double-Step Deep Learning Framework: Binary Cross-Entropy with Logits Loss (BCEWithLogitsLoss) for binary segmentation in Step One, ensuring accurate classification of burned and unburned regions, and Mean Squared Error (MSE Loss) for severity estimation in Step Two, optimizing the model to minimize the squared difference between predicted and actual severity levels. These hyperparameter choices were based on prior research and experimental validation, ensuring effective training and robust wildfire assessment.

4.3.5 Performance Metrics

In this thesis, the evaluation of model performance was conducted using Intersection over Union (IoU) and Root Mean Square Error (RMSE), which are also employed in the study of the DSF. These metrics provide a comprehensive assessment of the accuracy and reliability of both the binary segmentation and severity estimation models.

4.3.5.1 Intersection over Union (IoU)

Intersection over Union (IoU) is a widely used metric in segmentation tasks to quantify the overlap between the predicted and ground truth masks. IoU measures how well the predicted segmentation aligns with the actual burned area, making it a crucial indicator of the model's performance in wildfire detection. The IoU calculation is shown in Equation 4.1.

$$IoU = \frac{TP}{TP + FP + FN} \tag{4.1}$$

Where:

- TP (True Positives): Correctly predicted burned pixels.
- FP (False Positives): Incorrectly predicted burned pixels.

- FN (False Negatives): Incorrectly predicted unburned pixels.

A higher IoU value indicates better segmentation performance, as it signifies a greater degree of overlap between the predicted and actual burned area. This metric is particularly useful in evaluating the effectiveness of binary segmentation models, such as those based on U-Net variants and Mask R-CNN, in detecting wildfire-affected regions.

IoU is essential for measuring the accuracy of burned area detection in the binary segmentation step.

4.3.5.2 Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is utilized to assess the accuracy of severity prediction by quantifying the difference between predicted severity values and the actual ground truth severity levels. RMSE provides insight into how well the model can estimate severity levels based on burned area characteristics. The RMSE computation is presented in Equation 4.2.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{4.2}$$

Where:

- y_i is the actual severity level for a given burned area.
- $\hat{y}_i$ is the predicted severity level.
- N is the total number of observations.

A lower RMSE value indicates better model performance, as it suggests a smaller discrepancy between the predicted and actual severity levels. RMSE is particularly relevant in the second step of the DSF, where the severity regression model aims to predict the degree of fire damage within the burned regions.

5
EXPERIMENTAL RESULTS

In this study, the Double-Step Framework is used to assess the effectiveness of several variations of U-Net for the detection and segmentation of wildfire severity. In the experiments, U-Net, Attention U-Net, Nested U-Net, SegU-Net, and MultiResUNet models were tested. The first four models are integral components of the Double-Step Framework proposed in [12]. MultiResUNet, however, is included primarily for comparison purposes with minimal customization. Additionally, this study employs two distinct severity scales for the second step of the DSF: one with five levels and another with four levels. The 5-level classification allows for more detailed differentiation of severity, useful for ecological studies requiring precise segmentation of fire impacts. In contrast, the 4-level classification simplifies the analysis by addressing data imbalance and grouping levels 4 and 3.

5.1 Mask R-CNN

Mask R-CNN was evaluated independently from the Double-Step Framework to observe its performance on binary segmentation tasks. The model was trained using 3-channel satellite images as input and mask images for binary segmentation of burned areas. The results were poor, with IoU values across different test folds shown in Table 5.1, suggesting that the model failed to learn effectively. This could be attributed to the reliance on pretrained weights, which are likely optimized for general-purpose object detection tasks rather than wildfire-specific segmentation. The model's inability to generalize may also stem from the unique spectral and geographical characteristics of wildfire datasets, which require domain-specific training. Training the Mask R-CNN model from scratch, without using pretrained weights, could better adapt it to the wildfire dataset and improve its performance. Additionally, modifying the model to accept 12-channel Sentinel-2 images, which capture a broader range of spectral information, could further enhance its segmentation accuracy for wildfire datasets.

5.2 DSF with MultiResUNet

The performances are assessed using Intersection over Union (IOU) for binary predictions and Root Mean Square Error (RMSE) for regression tasks. IoU is a metric calculated by dividing the area of overlap between the predicted and ground truth regions by the area of their union. RMSE measures the average magnitude of the error between predicted and ground truth values, calculated as the square root of the mean of the squared differences. IoU is used to evaluate the binary segmentation performance in the first step of the framework, while RMSE is employed in the second step to assess the accuracy of severity estimation.

In the DSF model trainings, binary cross-entropy loss was used for the binary segmentation criterion, which measures the difference between predicted probabilities and actual binary labels (0 or 1) by calculating the negative log likelihood of the correct classification.

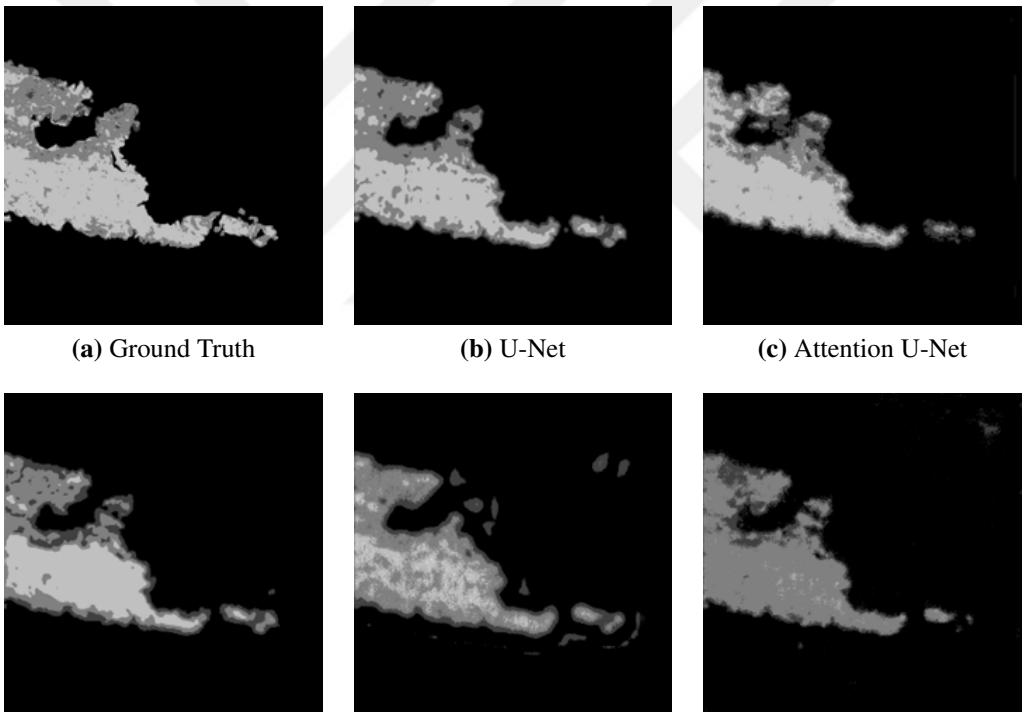
The first step of the DSF focuses on binary detection, distinguishing burned from unburned areas. IoU results, as shown in Table 5.1, include models within the DSF as well as an evaluation of Mask R-CNN as an external benchmark for binary segmentation. MultiResUNet demonstrated notable performance, achieving an IoU score of 0.84 on the Coral and Pink datasets. However, it is important to note that the model's predicted mask images often exhibited predicted lines close to the edge of the images, likely due to the minimal customization applied to the model. Attention U-Net also performed well in this step. It achieved an IoU of 0.82 in the Magenta dataset. Nested U-Net exhibited the most consistent performance overall with its highest IoU score of 0.85 on the Magenta and Pink datasets. The results of Step 1 suggest that MultiResUNet and Nested U-Net may be better suited for binary detection, as they achieved high IoU scores across different levels and datasets.

Table 5.1 Step 1 (Wildfire detection phase) results - binary IoU

Fold	**Double-Step Framework (12 Channels)**					**Single Step (3 Channels)**
	U-Net	**Attention U-Net**	**Nested U-Net**	**SegU-Net**	**MultiRes UNet**	**Mask R-CNN**
Coral	0.83	0.74	0.84	0.75	0.84	0.31
Cyan	0.66	0.68	0.68	0.59	0.66	0.27
Grey	0.48	0.71	0.78	0.70	0.64	0.22
Lime	0.56	0.60	0.57	0.57	0.58	0.38
Magenta	0.84	0.82	0.85	0.82	0.82	0.17
Pink	0.81	0.79	0.85	0.82	0.84	0.28
Purple	0.82	0.81	0.81	0.81	0.65	0.26

5.2.1 Level 4 and 5 Experiments

The second step of the Double-Step Framework is designed to perform a severity estimation on the detected burned areas, evaluating the models' ability to predict severity levels accurately. Table 5.2 shows the RMSE results for each dataset and their corresponding levels. MultiResUNet achieved the lowest RMSE value of 0.17 in the Lime dataset with the 4-level configuration. However, it is important to note that this RMSE result is at level 0, which represents unburned areas. Predicting unburned areas should be evaluated differently from other levels, since it largely reflects the model's binary detection accuracy rather than its capability for detailed severity classification. Nested U-Net displayed higher variability in RMSE values across different datasets, especially in the Coral dataset. The standard U-Net model maintained relatively stable RMSE results. In Figure 5.1 examples of predicted images are shown.



(a) Ground Truth **(b)** U-Net **(c)** Attention U-Net

(d) Nested U-Net **(e)** SegUNet **(f)** MultiResUNet

Figure 5.1 Example of predicted images

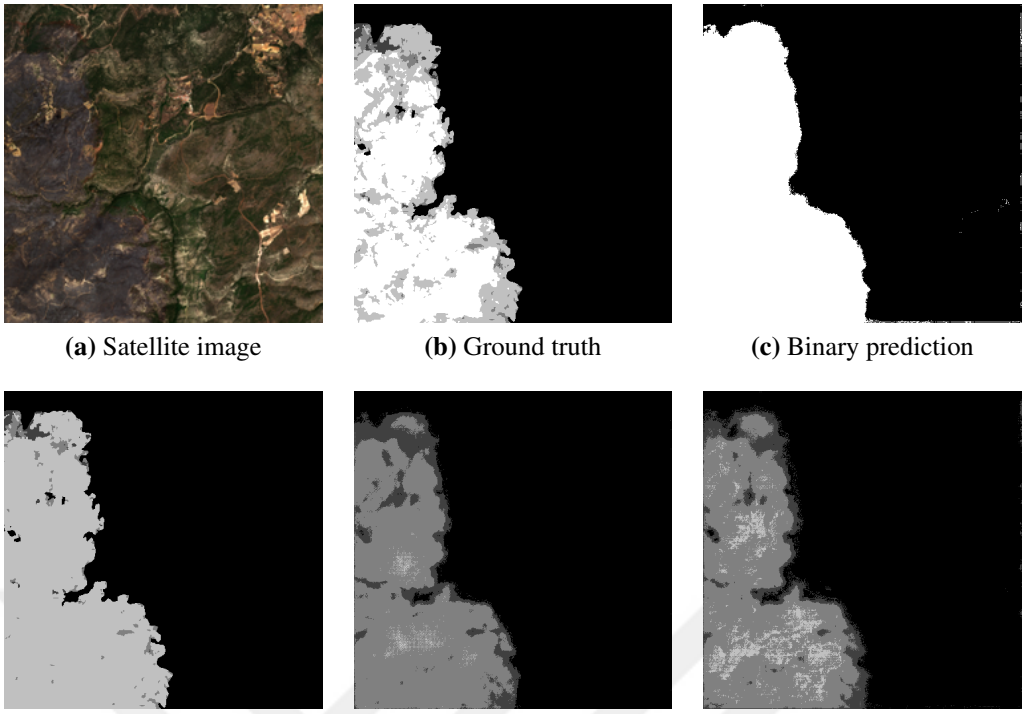
The experiments reveal that merging levels 4 and 3 into a 4-level classification consistently reduces RMSE values across most folds and models compared to the 5-level classification. This improvement is particularly notable in folds with significant data imbalance, such as Coral and Grey. For example, in the Coral fold, the RMSE for U-Net decreases from 0.79 (5-level) to 0.60 (4-level), and MultiResUNet exhibits similar trends with RMSE decreasing from 0.86 to 0.63. However, the Lime fold showed less consistent improvement, highlighting that the

Table 5.2 Step 2 (Wildfire severity classification) results - regression RMSE

Fold	Level	U-Net		Attention U-Net		Nested U-Net		SegU-Net		MultiRes UNet	
		5 Lvl	4 Lvl	5 Lvl	4 Lvl	5 Lvl	4 Lvl	5 Lvl	4 Lvl	5 Lvl	4 Lvl
Coral	0	0.71	0.61	0.66	0.49	0.56	0.54	0.42	0.44	0.48	0.47
	1	0.93	0.84	0.97	0.60	0.79	0.77	0.56	0.67	0.65	0.65
	2	1.04	0.60	1.01	0.86	0.88	0.53	0.83	0.79	0.54	0.63
	3	0.74	0.51	0.52	1.07	0.35	0.55	1.10	1.19	0.98	1.05
	4	0.83		1.79		1.52		2.48		2.42	
	ALL	0.79	0.60	0.86	0.69	0.72	0.55	0.90	0.70	0.86	0.63
Cyan	0	0.45	0.54	0.52	0.47	0.46	0.49	0.41	0.49	0.56	0.41
	1	1.05	1.08	1.09	1.02	1.04	0.93	0.81	0.65	1.17	0.90
	2	0.99	0.83	0.90	0.90	0.93	0.85	0.98	0.96	0.94	1.04
	3	0.71	1.36	0.68	1.44	0.74	1.39	1.14	1.80	0.67	1.62
	4	2.30		2.21		2.31		2.67		2.33	
	ALL	1.03	0.78	1.02	0.76	1.03	0.75	1.15	0.89	1.07	0.79
Grey	0	0.38	0.26	0.30	0.24	0.31	0.29	0.63	0.27	0.29	0.29
	1	1.14	0.95	1.15	0.72	0.91	0.88	0.70	0.75	0.89	0.93
	2	1.25	1.37	1.06	0.89	0.66	1.07	0.93	0.65	0.81	0.94
	3	1.23	1.47	1.16	1.55	1.21	1.27	1.55	1.44	1.47	1.30
	4	2.37		1.91		2.02		2.64		2.49	
	ALL	0.57	0.52	0.49	0.42	0.44	0.46	0.72	0.40	0.47	0.44
Lime	0	0.28	0.32	0.26	0.23	0.21	0.26	0.19	0.28	0.23	0.17
	1	0.76	0.63	0.75	0.73	0.71	0.65	0.82	0.80	0.77	0.78
	2	1.00	1.06	1.05	1.06	1.16	1.02	1.32	1.26	1.13	1.51
	3	1.60	1.48	1.85	1.61	2.03	1.48	2.30	1.90	1.95	2.27
	4	1.91		2.11		2.18		2.58		2.23	
	ALL	0.47	0.44	0.50	0.41	0.50	0.41	0.56	0.49	0.51	0.52
Magenta	0	0.38	0.42	0.39	0.45	0.34	0.39	0.34	0.26	0.29	0.48
	1	1.28	1.34	1.30	1.41	1.15	1.41	1.10	0.82	0.74	1.57
	2	0.84	0.84	0.96	0.80	0.79	0.84	0.84	0.76	0.76	0.87
	3	1.05	0.97	1.18	0.83	1.17	0.91	1.37	1.29	1.54	0.95
	4	1.24		1.27		1.42		1.72		2.12	
	ALL	0.63	0.62	0.67	0.61	0.63	0.60	0.70	0.57	0.74	0.68
Pink	0	0.32	0.23	0.28	0.25	0.23	0.26	0.26	0.24	0.18	0.21
	1	0.94	0.80	0.87	0.84	0.95	0.93	0.88	0.64	0.56	0.54
	2	0.56	0.79	0.92	0.80	0.66	0.60	0.78	0.55	0.93	0.70
	3	0.91	1.14	1.36	1.09	1.05	1.05	1.25	1.52	1.78	1.56
	4	1.69		2.55		1.66		1.78		2.26	
	ALL	0.45	0.48	0.57	0.48	0.45	0.46	0.52	0.57	0.65	0.58
Purple	0	0.20	0.21	0.28	0.20	0.20	0.26	0.24	0.17	0.22	0.34
	1	1.11	0.93	1.08	0.88	1.00	1.03	0.95	0.73	0.83	0.84
	2	0.54	0.51	0.71	0.44	0.48	0.57	0.36	0.50	0.69	0.56
	3	1.45	1.51	1.39	1.18	1.25	1.07	1.43	1.34	1.89	1.23
	4	1.74		1.67		1.85		2.03		2.24	
	ALL	0.57	0.41	0.59	0.43	0.60	0.43	0.65	0.46	0.71	0.51

effectiveness of this adjustment may vary depending on the dataset's severity level distribution. In Figure 5.2 examples of predicted images are shown.

In Table 5.3, the performance benchmarks of various models and methods employing the Double-Step Framework, including related works and the results from this study, are presented. The table includes metrics such as the Intersection over Union, Root Mean Squared Error for all levels (RMSE All), and Root Mean



(a) Satellite image **(b)** Ground truth **(c)** Binary prediction

(d) Level 4 mask representation **(e)** Level 5 severity representation **(f)** Level 4 severity representation

Figure 5.2 Example of MultiResUNet predictions based on levels

Squared Error for individual severity levels (RMSE Level). RMSE Level does not include Level 0 (unburned pixels), as unburned pixels dominate the dataset and could disproportionately influence the metric. By excluding Level 0, RMSE Level focuses on burned severity levels, providing a more meaningful comparison of the models' performance in assessing burned areas. The results allow for a comparative analysis of different approaches, highlighting the effectiveness of the DSF and related methods.

Table 5.3 Performance benchmark of related works

Model/Method	IoU	RMSE All	RMSE Level	Reference
DS-UNet	0.75	1.32	-	Monaco et al., 2021 [12]
DS-UNet++	0.74	1.41	-	
DS-SegU	0.65	1.66	-	
DS-AttU	0.72	1.38	-	
dNBR	-	0.91	0.48	Farasin et al., 2020 [11]
Single U-Net	-	0.95	0.35	
Parallel U-Net	-	0.97	0.35	
Double-Step U-Net	-	0.76	0.39	
MultiResUNet	0.84	0.44	0.54	This study

5.2.2 Combined Pre- and Post-Fire Imagery

In the binary burned area detection task, incorporating both pre-fire and post-fire Sentinel-2 images as a 24-channel input yielded improvements in several folds compared to using only post-fire images. Notably, the Cyan, Grey and Purple folds experienced IoU increases: from 0.66 to 0.73, 0.64 to 0.75, and 0.65 to 0.78, respectively. These gains suggest that the inclusion of temporal information enhances the model's ability to distinguish burned regions from unburned surroundings. The Coral fold maintained similar levels of performance. However, a slight decrease was observed in the Magenta and Pink folds. A more significant drop occurred in the Lime fold (from 0.58 to 0.48), which may be attributed to the complex terrain characteristics. Overall, the results demonstrate that using both pre-fire and post-fire imagery can enhance burned area segmentation performance.

The results for severity estimation reveal improvements when both pre-fire and post-fire images are combined as a 24-channel input. Several folds demonstrated notable gains in performance, such as the Coral and Pink folds. Similarly, Purple and Magenta experienced slight improvements in RMSE. The Lime fold showed a marginal gain (from 0.51 to 0.54). The Grey fold remained relatively stable. The Cyan fold had a slight degradation from 1.07 to 1.01, though still within a comparable range. These results suggest that the 24-channel configuration generally enhances model performance. However, improvements are not uniform across folds, potentially due to factors such as terrain complexity, seasonal variations.

5.2.3 Decoder Batch Normalization Removed

The binary segmentation results reveal that removing batch normalization from the decoder blocks of the MultiResUNet architecture generally led to a decline in performance across most folds. Compared to the original model, folds such as Pink and Lime exhibited the most noticeable decreases in IoU, dropping by 0.16 and 0.11 points respectively. Smaller performance drops were also observed in the Coral, Cyan, Grey, and Magenta folds. These results suggest that batch normalization in the decoder path plays an important role in stabilizing and improving performance. Interestingly, the Purple fold showed an improvement. This result suggests that in certain cases, removing batch normalization may help reduce overfitting and enhance generalization. Overall, the removal of decoder batch normalization appears to slightly reduce the effectiveness of binary burned area detection, with the exception of specific folds that may benefit from this architectural simplification.

The RMSE results for severity estimation reveal a mixed impact after removing

batch normalization from the decoder blocks of the MultiResUNet model. In some folds, such as Coral and Magenta, there was a clear improvement across most severity levels. Coral's overall RMSE decreased from 0.86 to 0.72. Magenta's RMSE dropped from 0.74 to 0.65. These two folds indicate better severity estimation with the modified architecture. The Purple fold also showed a moderate improvement (from 0.71 to 0.61). However, for other folds, the performance worsened. Lime saw a significant increase in RMSE, from 0.51 to 0.65. Cyan, Grey and Pink showed relatively stable performance, with only minor fluctuations in RMSE. Overall, these results suggest that removing batch normalization in the decoder introduces trade-offs: it may help reduce over-smoothing and improve predictions in some folds, while decreasing generalization or stability in others. The effect appears to be fold-dependent and possibly influenced by factors like data complexity, regional variability, or severity distribution. The results of both the batch normalization removed and the 24-channel input experiments are summarized in Table 5.4.

Table 5.4 MultiResUNet results with architectural variations

Fold	IoU			RMSE			
	12 Chnl	24 Chnl	Batch Norm Removed	Level	12 Chnl	24 Chnl	Batch Norm Removed
Coral	0.84	0.84	0.82	0	0.48	0.31	0.50
				1	0.65	0.94	0.57
				2	0.54	0.64	0.52
				3	0.98	0.31	0.72
				4	2.42	1.10	1.85
				ALL	0.86	0.72	0.72
Cyan	0.66	0.73	0.62	0	0.56	0.92	0.40
				1	1.17	0.94	0.85
				2	0.94	0.89	1.08
				3	0.67	0.92	0.98
				4	2.33	2.36	2.52
				ALL	1.07	1.01	1.10
Grey	0.64	0.75	0.59	0	0.29	1.49	0.31
				1	0.89	0.89	1.00
				2	0.81	0.78	0.96
				3	1.47	1.49	1.34
				4	2.49	2.07	2.52
				ALL	0.47	0.44	0.50
Lime	0.58	0.48	0.47	0	0.23	0.22	0.10
				1	0.77	0.70	0.84
				2	1.13	1.20	1.63
				3	1.95	2.10	2.67
				4	2.23	2.50	3.22
				ALL	0.51	0.54	0.65
Magenta	0.82	0.79	0.74	0	0.29	0.40	0.39
				1	0.70	1.21	1.45
				2	0.76	0.81	0.97
				3	1.54	1.21	1.07
				4	2.12	1.66	1.15
				ALL	0.74	0.71	0.65
Pink	0.84	0.82	0.68	0	0.18	0.20	0.29
				1	0.56	0.65	0.87
				2	0.93	0.67	0.96
				3	1.78	1.36	1.26
				4	2.26	2.38	2.58
				ALL	0.65	0.52	0.64
Purple	0.65	0.78	0.74	0	0.22	0.15	0.26
				1	0.83	0.92	1.20
				2	0.69	0.69	0.47
				3	1.89	1.60	1.49
				4	2.24	2.07	1.82
				ALL	0.71	0.64	0.61

6
CONCLUSION

This paper presents the implementation and evaluation of a Double-Step Deep Learning Framework for the detection and severity assessment of burned areas using Sentinel-2 satellite imagery. The Double-Step Framework leverages various U-Net model variations like Attention U-Net, Nested U-Net, SegU-Net, and MultiResUNet to assess wildfire impact across multiple severity levels. The first step in this framework successfully utilizes binary detection to distinguish burned from unburned regions, with MultiResUNet and Nested U-Net demonstrating strong performance in terms of IoU. Although MultiResUNet achieved high accuracy in the Coral dataset, predicted mask images exhibited lines close to the edge of the image, suggesting potential limitations due to the model's minimal customization.

The results of the 4-level and 5-level severity classifications demonstrate that merging levels 4 and 3 helps address data imbalance and reduces RMSE for underrepresented high-severity categories. This adjustment simplifies the classification task and often leads to better overall model performance, as seen in the lower RMSE values across most folds and models. However, the variability in performance across certain folds, such as Lime, underscores the need for further investigation into the impact of severity level distribution on model generalization.

The Mask R-CNN model was evaluated independently to benchmark its binary segmentation performance. Its results were significantly poor, likely due to the reliance on pretrained weights and the use of 3-channel input images instead of the full spectral range provided by Sentinel-2. Enhancing Mask R-CNN through domain-specific training and modifying its architecture to accept 12-channel Sentinel-2 imagery could improve its ability to process wildfire datasets effectively. Incorporating Mask R-CNN into the severity estimation step of the DSF represents a promising area for future exploration.

The experimental results from the second step show that Nested U-Net showed

greater variability in RMSE across different folds, with the Coral dataset posing particular challenges. The Pink fold achieved the best overall performance, suggesting that the distribution of severity levels in the dataset can impact the accuracy of the model. By contrast, the Lime fold exhibited the poorest detection results, likely due to the unique geographical characteristics of volcanic areas, which may introduce greater complexity in accurately detecting burned regions.

Even though MultiResUNet demonstrated good results for certain folds, such as achieving the lowest RMSE in the Lime dataset with the 4-level configuration, its overall performance in severity estimation was inconsistent across folds. This variability highlights a lack of generalizability in handling severity estimation tasks, particularly in datasets with complex geographical or spectral characteristics. Given these observations, MultiResUNet may not be the most suitable choice for severity estimation within the Double-Step Framework, despite its strong performance in binary segmentation.

This study demonstrates the potential of integrating multi-sensor deep learning models, leveraging the rich spectral information provided by Sentinel-2 imagery, for wildfire detection and severity assessment. Among the models tested, MultiResUNet with minimal customization showed similar results to other U-Net variants, particularly for burned area detection. This suggests that the first step of the Double-Step Framework could be optimized specifically for U-Net-based architectures. Furthermore, a potential area of future research could involve leveraging different model families, such as applying U-Net-based models for binary detection and architectures like Mask R-CNN for severity estimation. Additionally, employing more balanced satellite datasets could enhance the Double-Step Framework's effectiveness across a wider range of geographical features and severity levels.

In addition to the core experiments, two architectural variations of the MultiResUNet model were explored to assess their impact on performance. First, pre-fire and post-fire Sentinel-2 images were combined as a 24-channel input to incorporate temporal information. This approach led to improved binary segmentation performance in several folds. However, regression results were mixed. This can suggest that temporal stacking can be beneficial but remains sensitive to fold-specific characteristics such as terrain complexity. Second, an ablation experiment was conducted by removing batch normalization layers from the decoder blocks. This modification slightly improved regression performance in folds such as Coral, Magenta, and Purple. In terms of binary segmentation, removing batch normalization generally resulted in worse performance across most

folds, with Purple being the only exception showing improvement. These outcomes highlight that decoder normalization plays a nuanced role in severity estimation and that even minor architectural changes can differentially affect model generalization across diverse regions.

6.1 Future Works

While the Double-Step Framework has demonstrated promising results in burned area detection and severity assessment, several areas of improvement remain for future research.

One key direction is the enhancement of Mask R-CNN for binary segmentation. The current evaluation revealed that its performance was significantly limited due to its reliance on pretrained weights and the use of only 3-channel input images instead of the full 12-channel Sentinel-2 spectral range. Future work should explore domain-specific training and architectural modifications that allow Mask R-CNN to effectively process multi-spectral wildfire datasets. Furthermore, integrating Mask R-CNN into the severity estimation step of the framework represents a promising research avenue, as its object detection capabilities could improve high-severity category classification.

Another aspect for improvement involves severity level classification. The results indicate that merging severity levels 4 and 3 helps mitigate class imbalance and enhances model performance. However, further research is needed to explore alternative severity level distributions, as the Lime fold's performance variability suggests that different environmental conditions may impact model generalization. Expanding the dataset with balanced severity level distributions across diverse landscapes could lead to more robust severity assessment.

Future work could also investigate the optimization of U-Net-based models for binary detection. While MultiResUNet demonstrated strong results, it exhibited mask boundary inconsistencies, indicating potential limitations in its current form. Fine-tuning hyperparameters, introducing post-processing techniques, or experimenting with hybrid architectures that combine U-Net-based segmentation with other deep learning methods could address these issues. Additionally, future research should investigate why MultiResUNet struggles with consistency in severity estimation across different datasets. Enhancements to its architecture or loss functions may address the variability in its performance. For example, integrating weighted RMSE or other custom loss functions could help mitigate the impact of imbalanced severity levels.

Lastly, the application of different deep learning model families should be considered. Future studies could explore hybrid architectures, where U-Net-based models remain optimized for binary detection while architectures like Mask R-CNN or transformer-based networks are employed for severity estimation. Investigating these alternatives could provide further insights into improving the Double-Step Framework's adaptability across varying wildfire scenarios.

By addressing these future directions, the Double-Step Framework can be further refined to enhance its effectiveness in wildfire area segmentation and severity classification across diverse geographical and environmental conditions.

REFERENCES

[1] L. Colomba *et al.*, "A dataset for burned area delineation and severity estimation from satellite imagery," *Proceedings of the 31st ACM International Conference on Information amp; Knowledge Management*, ser. CIKM '22, Atlanta, GA, USA: Association for Computing Machinery, 2022, pp. 3893–3897, isbn: 9781450392365. doi: 10.1145/3511808.3557528. [Online]. Available: https://doi.org/10.1145/3511808.3557528.

[2] G. Navarro, I. Caballero, G. Silva, P.-C. Parra, Á. Vázquez, R. Caldeira, "Evaluation of forest fire on madeira island using sentinel-2a msi imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 58, pp. 97–106, Jun. 2017. doi: 10.1016/j.jag.2017.02.003.

[3] D. R. Cambrin, L. Colomba, P. Garza, "Cabuar: California burned areas dataset for delineation [software and data sets]," *IEEE Geoscience and Remote Sensing Magazine*, vol. 11, no. 3, pp. 106–113, Sep. 2023, issn: 2373-7468. doi: 10.1109/mgrs.2023.3292467. [Online]. Available: http://dx.doi.org/10.1109/MGRS.2023.3292467.

[4] A. Shamsoshoara, F. Afghah, A. Razi, L. Zheng, P. Z. Fulé, E. Blasch, *Aerial imagery pile burn detection using deep learning: The flame dataset*, 2020. arXiv: 2012.14036 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2012.14036.

[5] S. Shen, S. Seneviratne, X. Wanyan, M. Kirley, *Firerisk: A remote sensing dataset for fire risk assessment with benchmarks using supervised and self-supervised learning*, 2023. arXiv: 2303.07035 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2303.07035.

[6] I. El-Madafri, M. Peña, N. Olmedo-Torre, "The wildfire dataset: Enhancing deep learning-based forest fire detection with a diverse evolving open-source dataset focused on data representativeness and a novel multi-task learning approach," *Forests*, vol. 14, no. 9, 2023, issn: 1999-4907. doi: 10.3390/f14091697. [Online]. Available: https://www.mdpi.com/1999-4907/14/9/1697.

[7] D. Sykas, D. Zografakis, K. Demestichas, C. Costopoulou, P. Kosmidis, "Eo4wildfires: An earth observation multi-sensor, time-series machine-learning-ready benchmark dataset for wildfire impact prediction," Sep. 2023, p. 14. doi: 10.1117/12.2680777.

[8] G. H. de Almeida Pereira, A. M. Fusioka, B. T. Nassu, R. Minetto, "Active fire detection in landsat-8 imagery: A large-scale dataset and a deep-learning study," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178,

pp. 171–186, Aug. 2021, issn: 0924-2716. doi: 10.1016/j.isprsjprs.2021.06.002. [Online]. Available: http://dx.doi.org/10.1016/j.isprsjprs.2021.06.002.

[9] M. Sdraka *et al.*, *Floga: A machine learning ready dataset, a benchmark and a novel deep learning model for burnt area mapping with sentinel-2*, 2023. arXiv: 2311.03339 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2311.03339.

[10] P. Zhang, X. Hu, Y. Ban, "Wildfire-s1s2-canada: A large-scale sentinel-1/2 wildfire burned area mapping dataset based on the 2017–2019 wildfires in canada," *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 2022, pp. 7954–7957. doi: 10.1109/IGARSS46834.2022.9884275.

[11] A. Farasin, L. Colomba, P. Garza, "Double-step u-net: A deep learning-based approach for the estimation of wildfire damage severity through sentinel-2 satellite data," *Applied Sciences*, vol. 10, no. 12, 2020, issn: 2076-3417. doi: 10.3390/app10124332. [Online]. Available: https://www.mdpi.com/2076-3417/10/12/4332.

[12] S. Monaco *et al.*, "Attention to fires: Multi-channel deep learning models for wildfire severity prediction," *Applied Sciences*, vol. 11, no. 22, p. 11 060, 2021.

[13] N. Ibtehaz, M. S. Rahman, "Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation," *Neural Networks*, vol. 121, pp. 74–87, 2020.

[14] Y. Han, C. Zheng, X. Liu, Y. Tian, Z. Dong, "Burned area and burn severity mapping with a transformer-based change detection model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 13 866–13 880, 2024. doi: 10.1109/JSTARS.2024.3435857.

[15] M. M. Pinto, R. Libonati, R. M. Trigo, I. F. Trigo, C. C. DaCamara, "A deep learning approach for mapping and dating burned areas using temporal sequences of satellite images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 160, pp. 260–274, 2020, issn: 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2019.12.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271619303089.

[16] M. Li *et al.*, "Early forest fire segmentation based on deep learning," *2021 CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes (SAFEPROCESS)*, 2021, pp. 1–5. doi: 10.1109/SAFEPROCESS52771.2021.9693660.

[17] F. Khennou, M. Akhloufi, "Improving wildland fire spread prediction using deep u-nets," *Science of Remote Sensing*, vol. 8, p. 100 101, Sep. 2023. doi: 10.1016/j.srs.2023.100101.

[18] Z. Wang *et al.*, "Semantic segmentation and analysis on sensitive parameters of forest fire smoke using smoke-unet and landsat-8 imagery," *Remote Sensing*, vol. 14, no. 1, 2022, issn: 2072-4292. doi: 10.3390/rs14010045. [Online]. Available: https://www.mdpi.com/2072-4292/14/1/45.

[19] Orman Genel Müdürlüğü, *Orman yangınları eğitim sunuları*, Accessed: 2025-01-30, n.d. [Online]. Available: https://www.ogm.gov.tr/tr/e-kutuphane/kitaplik/orman-yanginlari-egitim-sunulari.

[20] General Directorate of Forestry (OGM), *Resmi istatistikler - orman yangınları*, Accessed: 2025-01-30, 2024. [Online]. Available: https://www.ogm.gov.tr/tr/e-kutuphane/resmi-istatistikler.

[21] Joint Research Centre, *Wildfires in the eu: 2022 was the second worst year, warning of changing climate*, Accessed: 2025-01-30, 2023. [Online]. Available: https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/wildfires-eu-2022-was-second-worst-year-warning-changing-climate-2023-11-22_en.

[22] European Environment Agency, "Forest fires in europe," Tech. Rep., 2021, Accessed: 2025-01-30. [Online]. Available: https://www.eea.europa.eu/en/analysis/indicators/forest-fires-in-europe.

[23] National Interagency Fire Center, *Wildfires in 2024*, https://www.nifc.gov/fire-information/statistics/wildfires, Accessed: 2025-01-31, 2024.

[24] General Directorate of Forestry (OGM), *Yamanlar'da yangından zarar gören orman alanları ağaçlandırılıyor*, Accessed: 2025-01-30, 2024. [Online]. Available: https://www.ogm.gov.tr/tr/haberler/yamanlar%E2%80%99da-yangindan-zarar-goren-orman-alanlari-agaclandiriliyor.

[25] National Interagency Fire Center (NIFC), *Burned area emergency response (baer) program*, Accessed: 2025-01-30, 2024. [Online]. Available: https://www.nifc.gov/programs/post-fire-recovery.

[26] European Soil Data Centre (ESDAC), *Post-fire soil erosion in europe*, Accessed: 2025-01-30, 2024. [Online]. Available: https://esdac.jrc.ec.europa.eu/themes/post-fire-soil-erosion.

[27] O. Ronneberger, P. Fischer, T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015. arXiv: 1505.04597 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1505.04597.

[28] O. Oktay *et al.*, *Attention u-net: Learning where to look for the pancreas*, 2018. arXiv: 1804.03999 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1804.03999.

[29] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, D. Stoyanov *et al.*, Eds., Cham: Springer International Publishing, 2018, pp. 3–11, isbn: 978-3-030-00889-5.

[30] U. Kamal, T. I. Tonmoy, S. Das, M. K. Hasan, "Automatic traffic sign detection and recognition using segu-net and a modified tversky loss function with l1-constraint," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1467–1479, 2020. doi: 10.1109/TITS.2019.2911727.

[31] K. He, G. Gkioxari, P. Dollár, R. Girshick, *Mask r-cnn*, 2018. arXiv: 1703.06870 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1703.06870.

[32] S. Ren, K. He, R. Girshick, J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, 2016. arXiv: 1506.01497 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1506.01497.

[33] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

PUBLICATIONS FROM THE THESIS

Papers

1. M. M. Yurdakul, B. Bayram, T. Bakırman, and H. O. İlhan, "A comparative study of double-step deep learning framework for burned area identification and severity assessment in wildfires," *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi,* vol. 28, no. 1, pp. 513–523, 2025, doi: 10.17780/ksujes.1601614.