Bachelor of Science in Computer Science & Engineering

Ensemble of Deep Learning methods for Precise Water Bodies Segmentation in Remote Sensing Imagery

by

Abu Shaleh Md.Toha

ID: 1804031

Department of Computer Science & Engineering

Chittagong University of Engineering & Technology (CUET)

Chattogram-4349, Bangladesh.

June, 2024

Ensemble of Deep Learning methods for Precise Water Bodies Segmentation in Remote Sensing Imagery

Submitted in partial fulfilment of the requirements for
Degree of Bachelor of Science
in Computer Science & Engineering

by

Abu Shaleh Md.Toha
ID: 1804031

Supervised by

Md.Obaidur Rahman
Associate Professor
Department of Computer Science & Engineering

Chittagong University of Engineering & Technology (CUET)
Chattogram-4349, Bangladesh.

The thesis titled '**Ensemble of Deep Learning methods for Precise Water Bodies Segmentation in Remote Sensing Imagery**' submitted by ID: 1804031, Session 2021-2022 has been accepted as satisfactory in fulfilment of the requirement for the degree of Bachelor of Science in Computer Science & Engineering to be awarded by the Chittagong University of Engineering & Technology (CUET).

Board of Examiners

_____ Chairman

Md.Obaidur Rahman
Associate Professor
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

_____ Member (Ex-Officio)

Dr. Abu Hasnat Mohammad Ashfak Habib
Professor & Head
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

_____ Member (External)

Dr. Kaushik Deb
Professor
Department of Computer Science & Engineering
Chittagong University of Engineering & Technology (CUET)

Declaration of Originality

This is to certify that I am the sole author of this thesis and that neither any part of this thesis nor the whole of the thesis has been submitted for a degree to any other institution.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. I am also aware that if any infringement of anyone's copyright is found, whether intentional or otherwise, I may be subject to legal and disciplinary action determined by Dept. of CSE, CUET.

I hereby assign every rights in the copyright of this thesis work to Dept. of CSE, CUET, who shall be the owner of the copyright of this work and any reproduction or use in any form or by any means whatsoever is prohibited without the consent of Dept. of CSE, CUET.

Signature of the candidate

Date:

Acknowledgements

First of all, all praise to the creator, I successfully completed writing my thesis. I received a lot of help and advice from a certain number of people to whom I will be grateful for the rest of my life. Without their support, I could not have finished my thesis. Firstly, I would like to express my gratitude and appreciation to my supervisor, Md.Obaidur Rahman, Associate Professor, Department of CSE, CUET, for all of his assistance and support throughout my thesis. From the very beginning, he guided me on how to progress with my thesis work. He listened and tried to solve all the problems I faced, even if they were the slightest. It helped me to overcome all of the difficulties I faced during my thesis work which I could have never done otherwise. My gratitude also leads to the defense members of the committee for their constructive criticism, suggestions for improvement, and approval of my efforts. I would also like to thank my wife and my sisters for supporting and encouraging me throughout this entire period.

Abstract

The segmentation of water bodies in very high-resolution satellite and aerial images is a crucial task in remote sensing, essential for applications such as environmental monitoring, urban planning, and disaster management. Traditional methods often struggle to cope with the intricacies and high variability of these high-resolution images. This thesis addresses these challenges by leveraging the DeepLabv3+ model, a architecture for semantic segmentation, enhanced through the integration of an ensemble of feature extraction backbones. Our approach employs ResNet50V2, MobileNetV3, and EfficientNetV2 as backbones in the encoder part of the DeepLabv3+ model, utilizing a weighted averaging ensemble technique to combine their strengths. This ensemble method aims to improve the feature extraction process and enhance the overall segmentation performance. The backbones were selected for their complementary characteristics: ResNet50V2 for its robust feature extraction, MobileNetV3 for its computational efficiency, and EfficientNetV2 for its advanced scaling capabilities. The results demonstrate significant improvements in key evaluation metrics, with the ensemble model achieving approximately 96% accuracy and a mean Intersection over Union (mean IoU) of 88%. Additionally, metrics such as the F1-score, precision, and recall also showed marked enhancements, confirming the effectiveness of the weighted averaging technique in aggregating the outputs of the different backbones. Our research provides a comprehensive analysis of the benefits of using an ensemble of ResNet50V2, MobileNetV3, and EfficientNetV2 backbones in the DeepLabv3+ model for the segmentation of water bodies in high-resolution remote sensing images. The promising results underscore the potential of ensemble techniques in improving segmentation performance.

Keywords: Water Bodies segmentation, DeepLabV3+, ResNet50V2, remote sensing imagery, Semantic segmentation, Atrous convolution, EfficientNetV2, MobileNetV3.

Table of Contents

Acknowledgements **iii**

Abstract **iv**

List of Figures **vii**

List of Tables **viii**

1 Introduction **1**

1.1 Introduction 1
1.2 Framework/Design Overview 3
1.3 Difficulties 4
1.4 Applications 5
1.5 Objectives 6
1.6 Motivation 6
1.7 Contribution of the thesis 7
1.8 Thesis Organization 8
1.9 Conclusion 9

2 Literature Review **10**

2.1 Introduction 10
2.2 Related Literature Review 10
2.3 Conclusion 13

3 Methodology **14**

3.1 Introduction 14
3.2 Overview of the Methodology 14
3.2.1 Dataset Preprocessing 15
3.2.2 DCNNs as Backbone with DeepLabV3+ 17
3.2.2.1 Key Components of the Deeplabv3+ Model . . . 18
3.2.3 ResNet50V2 Backbone Architecture 19
3.2.3.1 Key Components of ResNet50V2 20
3.2.3.2 Implementation of ResNet50 as a Feature Extractor 21
3.2.4 MobileNetV3 Backbone Architecture 21
3.2.4.1 Key Components of MobileNetV3 22

3.2.4.2 Implementation of MobileNetV3 as Feature Extractor . . . 22
3.2.5 EfficientNetV2 Backbone Architecture . . . 23
3.2.5.1 Key Components of EfficientNetV2 . . . 23
3.2.5.2 Implementation of EfficientNetV2 as Feature Extractor . . . 23
3.2.6 Training Details . . . 24
3.2.7 Evaluation of the Model . . . 25
3.3 Conclusion . . . 26

4 Results and Discussions **27**
4.1 Introduction . . . 27
4.2 Dataset Description . . . 27
4.3 Experimental Setup . . . 28
4.4 Evaluation of Performance . . . 28
4.4.1 Evaluation Metrics . . . 28
4.4.2 Comparison of different pre-trained model . . . 32
4.4.3 Result Analysis of the Ensemble model . . . 33
4.4.4 Performance Evaluation in Relation to Previous Literature 35
4.5 Impact Analysis . . . 36
4.5.1 Social Environmental Impact . . . 37
4.5.2 Economic impact . . . 38
4.5.3 Ethical Impact . . . 39
4.6 Conclusion . . . 41

5 Conclusion **42**
5.1 Conclusion . . . 42
5.2 Future Work . . . 43

List of Figures

1.1 Proposed Framework 3

3.1 Proposed Methodology 15
3.2 Resizing from Original image 16
3.3 Proposed DeepLabv3+ Model with DCNN as Backbone 18
3.4 ResNet50 Model Architecture 20
3.5 Result of the input images after evaluation 26

4.1 Graphical representation of IoU 31
4.2 Accuracy Curve 33
4.3 Loss Curve 34
4.4 Precision Curve 34
4.5 F-1 Score Curve 35
4.6 mIoU Curve 35

List of Tables

4.1 Different Pre-trained Models as Backbone for Water bodies Segmentation 32
4.2 Results by using Different Pre-trained Models as Backbone 32
4.3 Comparison among different existing methods and proposed method 36

Chapter 1

Introduction

1.1 Introduction

The accurate and efficient segmentation of water bodies in very high-resolution satellite and aerial images is critical for various applications in remote sensing, including environmental monitoring, urban planning, and disaster management.[1] Remote sensing technologies provide valuable tools for acquiring detailed spatial information over large areas, enabling precise analysis and decision-making. However, the complex nature of high-resolution imagery poses significant challenges for traditional image processing techniques, necessitating the adoption of advanced machine learning models for accurate and automated segmentation tasks.[2][3]

In recent years, deep learning has emerged as a powerful paradigm in image analysis, offering state-of-the-art solutions for semantic segmentation tasks[4]. Among the deep learning models, DeepLabv3+ has shown exceptional performance in segmenting objects with fine details and boundaries. This thesis focuses on enhancing the segmentation accuracy of water bodies by leveraging the DeepLabv3+ model and introducing a novel ensemble technique.

The primary objective of this research is to investigate the effectiveness of an ensemble approach using ResNet50V2, MobileNetV3, and EfficientNetV2 as backbones in the DeepLabv3+ architecture. By employing a weighted averaging ensemble method, we aim to capitalize on the unique strengths of each backbone: ResNet50V2 for its robust feature extraction capabilities, MobileNetV3 for its computational efficiency, and EfficientNetV2 for its superior scaling properties. This ensemble strategy is expected to improve the model's ability to capture

intricate details and variations in water region imagery, leading to enhanced segmentation accuracy and robustness. Accurate and timely identification of water bodies is a critical component of management. Traditional methods for water mapping, such as ground surveys and hydrological modeling, are often time-consuming, labor-intensive, and limited in spatial coverage[5]. Remote sensing technology, which involves acquiring and analyzing satellite or aerial imagery, has emerged as a powerful tool and mapping due to its ability to provide large-scale and real-time data[6]. Recent advancements in artificial intelligence [7], particularly deep learning, have further enhanced the potential of remote sensing for water bodies segmentation. Deep learning models, with their ability to automatically learn and extract complex features from raw data, have shown great promise in various image analysis tasks. Among these models, convolutional neural networks (CNNs) have been particularly successful in the field of image segmentation.

This thesis focuses on the application of DeepLabV3+, a state-of-the-art deep learning model, for water bodies segmentation from remote sensing imagery. DeepLabV3+ is a semantic segmentation model that utilizes atrous convolution and spatial pyramid pooling to capture multi-scale contextual information. The model employs ResNet-50 as its encoder for feature extraction, significantly enhancing its ability to produce accurate segmentation maps.[8]

The primary objective of this research is to develop and evaluate a robust segmentation model that can accurately delineate water affected areas from satellite images. A comprehensive dataset comprising multi-spectral satellite images and corresponding ground truth annotations was used to train and validate the model. The performance of DeepLabV3+ was assessed using metrics such as Intersection over Union (IoU), precision, recall, and F1 score.

The results of this study demonstrate that DeepLabV3+ achieves high accuracy in water bodies segmentation, outperforming traditional image processing techniques and other segmentation models. By leveraging advanced deep learning methodologies, this research contributes to the development of more effective flood response strategies and disaster management practices. Future work may

involve integrating additional data sources, such as topographic and meteorological data, to further enhance model accuracy and predictive capabilities.

1.2 Framework/Design Overview

The proposed methodology for water bodies segmentation using DeepLabV3+ is a comprehensive approach designed to accurately identify and delineate water areas within satellite or aerial images. It begins with the collection and preparation of the dataset, which involves gathering images along with corresponding ground truth masks, resizing the images, and encoding the masks for training purposes. The model selection phase focuses on choosing DeepLabV3+ as the semantic segmentation model, utilizing pre-trained ResNet50V2, EfficientNetV2, MobileNetV3 backbone to leverage transfer learning. The overview of all methodology steps is presented in this section and 1.1 represents the block diagram of the proposed model.

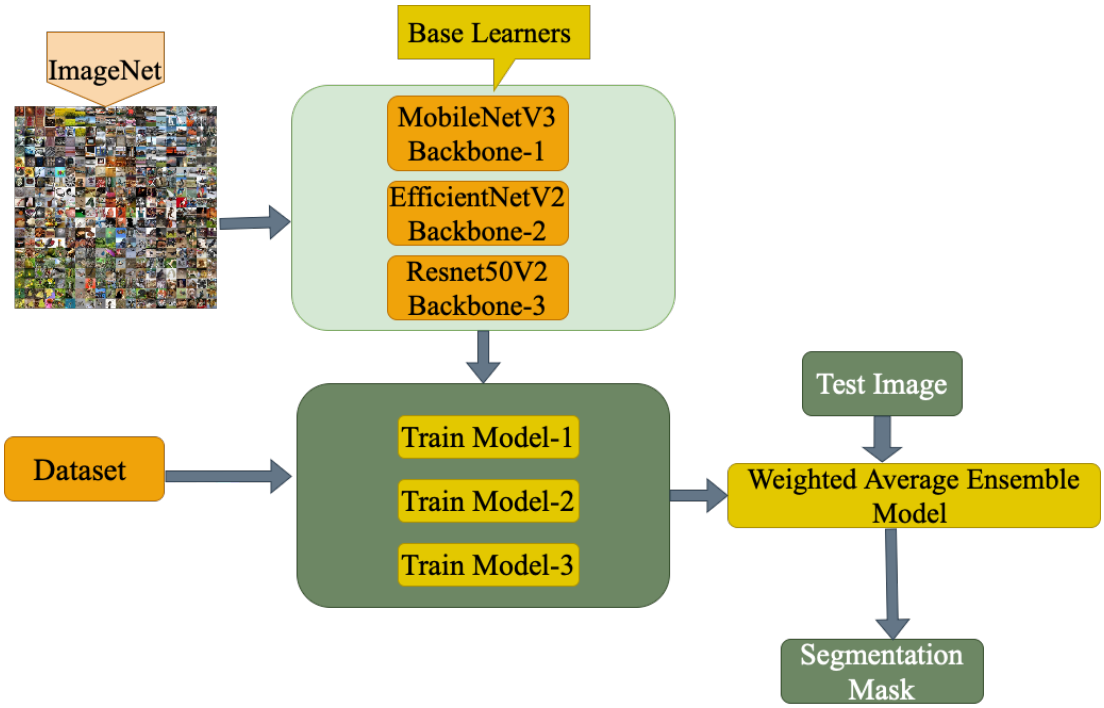


Figure 1.1: Proposed Framework

1.3 Difficulties

There are always some difficulties while working with deep learning methods and different dataset.Our thesis project on water bodies segmentation is crucial as it provides a realistic account of the challenges faced and how they were addressed. A list of some of the most common challenges is provided below:

1. **Data Collection :** Finding sufficient and high-quality satellite images, drone footage and normal images as well as their corresponding masks were challenging.

2. **Data Quality :** Differences in image resolution affected the consistency and accuracy of the segmentation results.

3. **Annotation and Labeling :** Ensuring consistent and accurate labeling across the dataset were difficult, especially with complex scenes.

4. **Finding the appropriate model :** Numerous candidate models with different parameters that can be adjusted in different ways to improve performance are available. Therefore, a large number of experiments must be conducted in order to identify the effective model.

5. **Model Development :** Fine-tuning hyperparameters for the optimal model performance was computationally intensive and required significant resources.

6. **Overfitting :** It's crucial to stop overfitting when the DNN model is being trained. It is necessary to use methods like cross-validation and regularization to make sure the model performs effectively when applied to new data.

7. **Training the model :** It requires a significant amount of time and machine configurations to train many models with varying parameters. Finding the ideal parameters and model for our architecture grew more difficult as a result.

1.4 Applications

- **Environmental Monitoring:** Accurate segmentation of water bodies can aid in monitoring water quality parameters such as turbidity, sedimentation, and pollution levels. This information is crucial for assessing environmental health and implementing remedial measures. Identifying and mapping water bodies helps in understanding aquatic habitats and ecosystems, supporting biodiversity conservation efforts.
- **Urban Planning and Infrastructure Management:** Precise delineation of water bodies assists in managing water resources, including allocation for drinking water supply, irrigation, and industrial use.
- **Flood Risk Assessment:** Mapping water bodies contributes to flood risk assessment and mitigation strategies, allowing urban planners to better plan infrastructure and develop early warning systems.
- **Disaster Management:** During natural disasters such as floods or hurricanes, accurate maps of water bodies facilitate emergency response operations, evacuation planning, and assessing the extent of damage. After disasters, monitoring changes in water bodies helps in assessing environmental impact, planning restoration efforts, and prioritizing resource allocation.
- **Agriculture and Land Use Planning:** Identifying water bodies aids in planning irrigation systems and optimizing agricultural practices, ensuring efficient water use in agriculture.
- **Infrastructure Development and Engineering:** Accurate mapping of water bodies informs decisions on infrastructure development projects such as dams, bridges, and roads, minimizing environmental impact and optimizing construction planning.
- **Hydrological Modeling:** Detailed water body segmentation supports hydrological modeling for flood forecasting, groundwater recharge estimation, and watershed management.
- **Remote Sensing and GIS Applications:** Water body segmentation is

essential for creating detailed and up-to-date maps used in GIS applications for various purposes, including navigation, tourism, and land use planning.

1.5 Objectives

The objectives are designed to address key challenges in water bodies segmentation and to develop a practical, accurate, and user-friendly model that can be utilized in various real-world applications.The key objective and possible outcomes of this work are mentioned below:

- To design and implement a robust model for accurately segmenting water areas in satellite or aerial imagery.
- To improve existing image processing techniques or develop new ones for better feature extraction and classification of water regions.
- To rigorously evaluate the performance of the segmentation model using various metrics, such as accuracy, precision, recall, and F1 score, Mean IoU.
- To test the segmentation model in real-world scenarios and assess its applicability for various use cases.
- To identify and analyze the challenges and limitations encountered during the development and implementation of the segmentation model.

1.6 Motivation

Accurate segmentation of water bodies supports various environmental monitoring applications, including assessing water quality, monitoring changes in aquatic habitats, and understanding the impacts of climate change on hydrological systems. These applications are crucial for informed decision-making in environmental management and policy.With the increasing availability and volume of satellite imagery, there is a growing demand for automated and efficient methods to analyze and interpret these vast datasets. Automated segmentation of water

bodies reduces manual effort, accelerates analysis processes, and improves the consistency and accuracy of results.

The factors that led us to motivate for this work :

- One of our main motivations was traditional water monitoring methods can be slow and resource-intensive. There is a growing need for advanced, automated tools that can quickly and accurately identify water region areas, helping authorities to respond more effectively to these events.
- Water bodies, such as lakes, rivers, and reservoirs, are crucial natural resources that support various ecosystems and human activities. Accurate mapping and monitoring of these water bodies are essential for effective water resource management, environmental conservation, and disaster preparedness.
- Advances in remote sensing technology and machine learning have created new opportunities for developing sophisticated water detection systems. Leveraging these technologies can enhance our ability to monitor and manage flood risks.
- Climate change is contributing to more frequent and severe flooding events. Developing robust water segmentation models is essential for preparing and adapting to these changing environmental conditions.
- Despite the advancements in deep learning for image segmentation, there may still be gaps in achieving optimal accuracy and efficiency, especially in the context of diverse and challenging environmental conditions. Your thesis aims to fill this gap by proposing and evaluating an innovative ensemble approach to enhance water body segmentation performance.

1.7 Contribution of the thesis

The contributions of this thesis are multifaceted, addressing both theoretical and practical aspects of water bodies segmentation. By developing a robust and accurate segmentation model, integrating multi-source data, and providing an

end-to-end solution for real-time water monitoring, this research makes significant advancements in the field. The primary contribution of this thesis is the following:

1. **Development of a Robust Segmentation Model :** This thesis develops a novel segmentation model using deep learning techniques. The model is capable of accurately identifying and delineating water areas from satellite or aerial imagery, offering improvements over existing methods in terms of accuracy and computational efficiency.

2. **Integration of Multi-Source Data :** The research integrates data from various sources, such as satellite imagery, weather data, and topographical information, to enhance the model's performance. This multi-source data fusion helps in achieving more reliable and comprehensive water mapping.

3. **Application of Advanced Machine Learning Techniques :** The architecture of the neural networks was fine-tuned by experimenting with different numbers of layers, filter sizes, and activation functions to find the optimal configuration for the segmentation.

1.8 Thesis Organization

The complete report is divided into five chapters. The report structure may be summarized as follows:

- **Chapter 1** provides a general description of research and discusses the obstacles, motivation, and application of the work.

- **Chapter 2** outlines and compares the technology employed in our works with some related works

- **Chapter 3** describes the technique and illustrates the system's implementation

- **Chapter 4** discusses the experimental results and analysis of the experiment.

- **Chapter 5** contains the conclusion of this work which includes an overall overview of our system and also mentions some limitations & future work.

1.9 Conclusion

We provided a thorough description of the design in this chapter, covering its conception, problems that were found, possible uses, and contributions to the world at large. We also discussed the work's larger context and significance, highlighting its relevance, goals, and possible ramifications.

Chapter 2

Literature Review

2.1 Introduction

Water bodies segmentation is a critical component in disaster management and mitigation, providing essential information for emergency response and urban planning. The advent of deep learning has significantly advanced the field, offering more accurate and efficient methods for water bodies detection and segmentation. This literature review explores the evolution of water bodies segmentation techniques, the impact of deep learning, and the integration of multi-source data to enhance model performance.

In this chapter, we will go through a comprehensive review of the recent research works in the field of water bodies segmentation, discuss their contribution and limitations with a comparative perspective.

2.2 Related Literature Review

The advent of deep learning revolutionized segmentation by leveraging neural networks, particularly Convolutional Neural Networks (CNNs), to automatically learn hierarchical features from data. These methods significantly outperformed traditional and machine learning approaches.Various researchers have contributed novel approaches to address the challenges in this field.

L. Meng et al. [8] proposed a method that a flood mapping model using HISEA-1 SAR imagery with an enhanced DeepLabv3+ framework. It employs MobileNetv2, smaller ASPP dilation rates, and extra upsampling layers, achieving superior accuracy and outperforming U-Net, SegNet, and standard DeepLabv3+.

Applied to 2021 floods in Henan and New Orleans, it showed significant inundation detection.The study highlights SAR's potential in disaster monitoring and anticipates more low-cost SAR satellites for improved emergency response.

C. Sazara et al. [9] proposes novel flood image classification methods using VGG-16 and logistic regression, achieving high performance. Flood area segmentation methods were compared, showing promising results for both superpixel-based and FCN approaches. FCN has room for improvement with more labeled data. Future work includes testing more advanced segmentation networks, extracting floodwater information like severity and depth, and enhancing the model for water reflection cases.

In [10], The study proposes an image-based method using passive monitoring cameras. It detects abnormal water level fluctuations, providing accurate monitoring images and water level data for disaster prevention in small urban areas. This approach supplements the current methods relying on in situ and remote sensing data, enhancing timely decision-making for reducing flood disasters.

In [11], The study focuses on robust water detection for UGV autonomous navigation, as traversing through deep water can damage UGV electronics and disrupt military missions. Previous work detected water bodies at mid to far range using sky reflections but faced challenges at close range. The study explores detecting water based on color variations from leading to trailing edges, using saturation and brightness changes. Software was developed to identify candidate water regions with low texture, evaluate color changes, and apply ellipse fitting for final water detection. This approach effectively identifies water bodies, aiding UGV navigation and mission safety.

In [12], The study presents an image segmentation method for aerial images, demonstrating its effectiveness in hydrological modeling. Although there are some errors, the results are comparable to hand-labeled data. The approach is general and can be improved further by considering shadows and increasing the number of classes for peak flow estimation.

In [13], The article introduces an Artificial Neural Network (ANN) approach for flash flood sensing using a custom-designed sensor with ultrasonic rangefinder and

infrared temperature sensors. ANNs accurately estimate temperature deviations affecting sound speed, outperforming other models. The method enables real-time water level monitoring in streets with high precision and robustness. Future work will explore rain detection using ultrasonic reflections and water presence detection for fault identification purposes. The approach significantly reduces power usage and bandwidth in wireless sensor networks.

In [14], The paper introduces a novel approach for detection of over water, focusing on visual features not previously explored in the literature. The method uses a probabilistic model to determine the flood region's position in images by analyzing color, contrast, entropy, and their dynamic changes across frames. Unlike complex feature extraction methods, this approach allows for fast processing, suitable for real-time detection and video retrieval in news content. Experimental results demonstrate the method's applicability.

In [15], The paper applies the Mask-R-CNN algorithm to detect and segment floodwater in urban, suburban, and natural scenes with high accuracy. The model achieves a good accuracy in floodwater detection as well as in segmentation. However, the coarse pixel resolution of the segmentation results is noted as an artifact. Future studies will explore floodwater depth prediction, improve detection speed, and address water reflection challenges to enhance segmentation accuracy.

In [16], The paper presents a model for water detection based on dynamic texture, enabling safe robot navigation in natural environments. The method utilizes entropy measurements from optical flow trajectories to detect water regions, even in less rippled areas. Future work will explore probabilistic models, spatio-temporal smoothing, and tracker stabilization techniques to enhance accuracy and robustness.

In [17], The study estimates water extent using simple color segmentation and depth analysis on crowd-sourced random images. This method provides valuable information for providing help to flood-affected areas by approximating the percentage of submerged area and floodwater volume from the images. Future work

aims to enhance accuracy by considering additional factors such as vehicles and human body parts relative to the water level to calculate the flood extent.

The paper [18] focuses on the increasing frequency and impact of water level, emphasizing the need for effective mitigation strategies. It presents a flood segmentation and visualization approach using the U-Net architecture with Sentinel-1 SAR satellite imagery. U-Net captures relevant features in SAR images, and the method includes data loading, preprocessing, flood inference, and visualization. The Sen1Floods11 dataset is used for training and validation. This study demonstrates the potential of deep learning for flood prediction and mapping.

The paper [19] highlights the inadequacy of ground-based segmentation models for aerial disaster imagery, especially during events like hurricanes. It evaluates real-time and non-real-time models on the FloodNet dataset from Hurricane Harvey, finding UNet-MobileNetV3 achieves 59.3% mIoU and PSPNet 79.7% mIoU. The study underscores the trade-off between accuracy and efficiency and the need for specialized aerial segmentation models for effective emergency response.

2.3 Conclusion

This chapter provides a comprehensive review of related works in the field of water bodies segmentation. Water segmentation reveals a transformative shift from traditional methods to deep learning approaches, particularly Convolutional Neural Networks (CNNs), which have demonstrated superior performance in accurately delineating extents. Integration of multi-source data, including optical, Synthetic Aperture Radar (SAR), and topographic information, has further enhanced segmentation accuracy by capturing diverse environmental characteristics. Despite notable advancements, challenges such as limited annotated data and computational constraints persist, necessitating innovative solutions for data collection and model refinement. The convergence of machine learning and remote sensing technologies offers promising avenues for improving water detection and mapping in real-time scenarios. Future research should focus on developing hybrid models that leverage the strengths of both traditional and deep learning techniques to address the complexities.

Chapter 3

Methodology

3.1 Introduction

In this chapter, we will jump into the methodology of our work. To begin, we will provide a comprehensive overview of the system. We will then proceed to explain the details in a step-by-step manner. Finally, we will showcase the implementation of our thesis.It involves several stages including data preparation, model architecture, training, and evaluation. Below is a detailed breakdown of the methodology employed in this project.

3.2 Overview of the Methodology

The methodology for the water bodies segmentation using DeepLabV3+ with weighted avearage ensemble method is structured into several critical steps. Initially, the system setup involved installing necessary libraries and dependencies and configuring the environment for reproducibility. The dataset was then prepared by defining essential configurations such as image size, batch size, number of classes, and augmentation factors, followed by splitting it into training, validation, and test sets. Data processing involved reading and preprocessing images and masks, with functions handling resizing, normalization, and augmentations to ensure the data was in the appropriate format for model training. TensorFlow datasets were created from the images and masks, applying preprocessing functions to make the data ready for the DeepLabV3+ model. This structured approach ensures the model can be effectively trained and evaluated for the task of water segmentation. Figure 3.1 depicts the overview of our proposed Methodology.

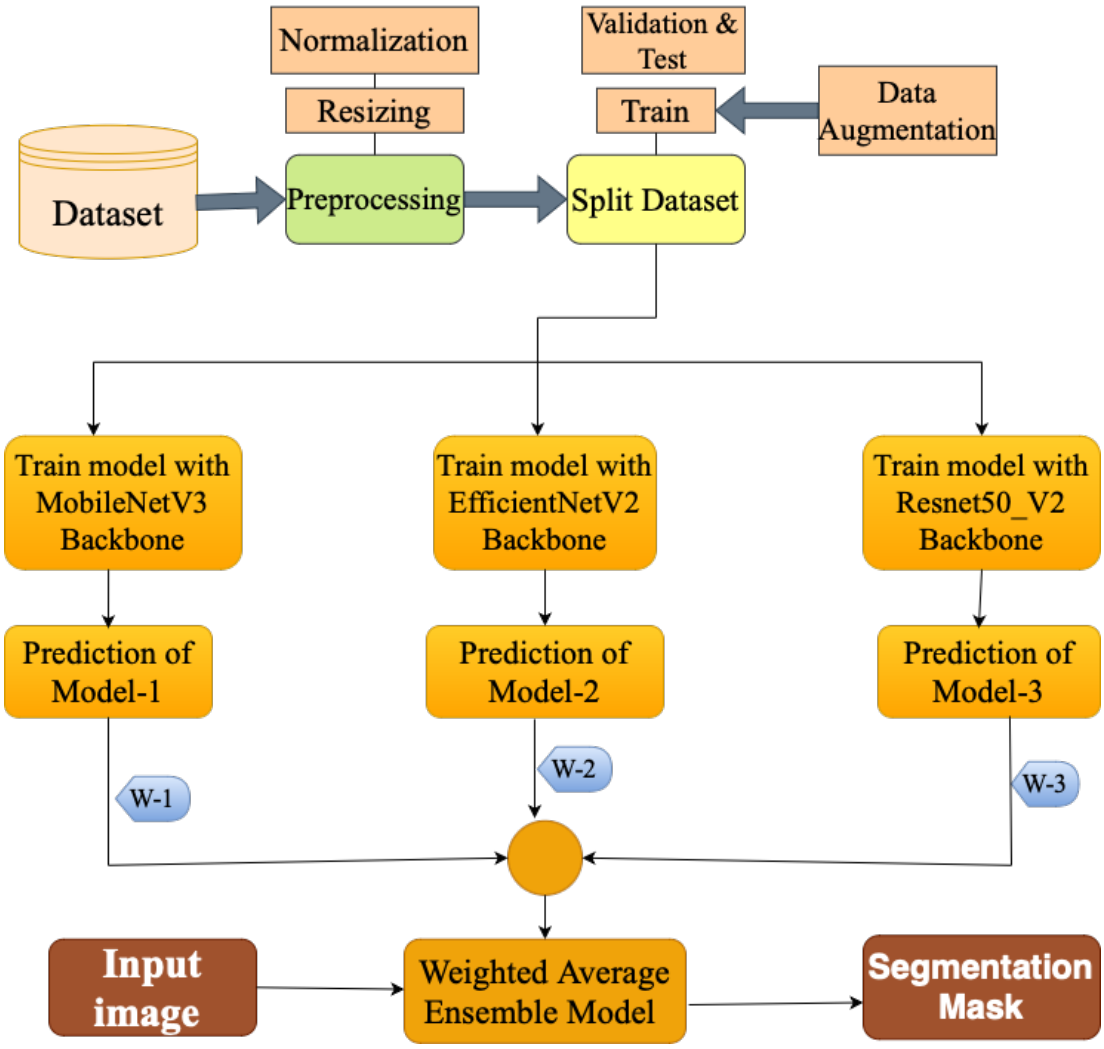


Figure 3.1: Proposed Methodology

3.2.1 Dataset Preprocessing

1. **Resizing:** In the context of preparing datasets for machine learning models, resizing is a critical preprocessing step. For this thesis, resizing involves adjusting the dimensions of the images and their corresponding segmentation masks to ensure uniformity and compatibility with the DeepLabV3+ model, which simplifies the training process. Below are the detailed steps and rationale for resizing the datasets:

    - **Uniform Image Size:** Machine learning models, especially Convolutional Neural Networks (CNNs) like DeepLabV3+, require inputs of consistent size. Resizing all images and their corresponding masks

to the same dimensions ensures that each input tensor has the same shape, which simplifies the training process and model architecture.

- **Memory and Computational Efficiency :** Resizing images to a smaller, fixed size (e.g. 256x256 pixels), which shows in 3.2 reduces the computational load and memory usage. This makes it feasible to train the model on available hardware, especially when dealing with large datasets.
- **Aspect Ratio Considerations:** While resizing, it's important to maintain the aspect ratio to avoid distortion. However, in this implementation, the images are uniformly resized, assuming that the original aspect ratios do not vary significantly, or that the impact of this distortion is minimal for the given task.

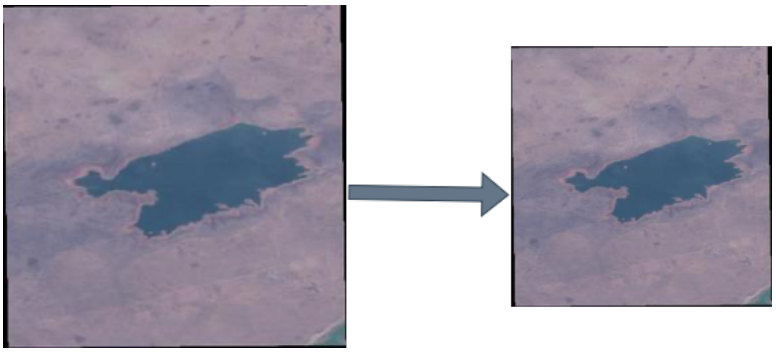


Figure 3.2: Resizing from Original image

2. **Normalization:** In our thesis, various normalization techniques are applied to both the input images and the corresponding masks. For the images, brightness and contrast adjustments are performed. These adjustments standardize the brightness and contrast levels across the dataset, helping the model to learn more robust features. Additionally, images are rescaled by clipping pixel values to the range of 0 to 255 and casting them to tf.float32, ensuring that the pixel values are in a standard range suitable for processing by the model.

   For the mask images, a binary mask creation technique is employed. This involves normalizing the pixel values by setting those greater than or equal

to a threshold value (200) to 1 and others to 0. This step is crucial for semantic segmentation tasks as it helps distinguish between the background and the target object, which in this case is the waterbody. By creating binary masks, the model can more effectively learn to segment the water areas.

3. **Image Augmentation:** Image augmentation is a technique used in data preprocessing to artificially increase the size and diversity of a dataset by applying various transformations to the original images. It helps improve the generalization and robustness of machine learning models by exposing them to a wider range of variations and scenarios.Five common types of image augmentation that have been used in this research are:

   - Horizontal Flip: The image is flipped horizontally along the verticalaxis, effectively mirroring it.
   - The image is flipped vertically along the horizontal axis.
   - The image is rotated 90 degrees in a clockwise direction.
   - The image is rotated 180 degrees in a clockwise direction.

   These augmentation techniques introduce variations in the dataset, enabling the model to learn features and patterns from different orientations and perspectives. By applying these transformations, the dataset becomes more diverse, allowing the model to better handle real-world scenarios with variations in orientation, lighting, and other facto

4. **Conversion to TensorFlow Dataset:** Converting the preprocessed images into a TensorFlow dataset involves transforming the image data into a format that TensorFlow can efficiently use for training machine learning models.

3.2.2 DCNNs as Backbone with DeepLabV3+

Utilized ResNet50V2, MobileNetV3, EfficientNetV2 as the backbone architecture for feature extraction, leveraging its pre-trained weights on the ImageNet dataset. Then employed the DeepLabV3+ architecture for semantic segmentation, which

incorporates atrous convolution and encoder-decoder structure for capturing contextual information and precise segmentation.

The feature maps extracted by these Deep Convolutional Neural Networks (DCNNs) which are used as the backbone network ,serve as input to the DeepLabV3+ architecture. These feature maps capture rich semantic information from the input images, which is essential for accurate segmentation. The Below figure 3.3 shows the graphical representation of the DeepLabv3+ model with DCNN backbone.

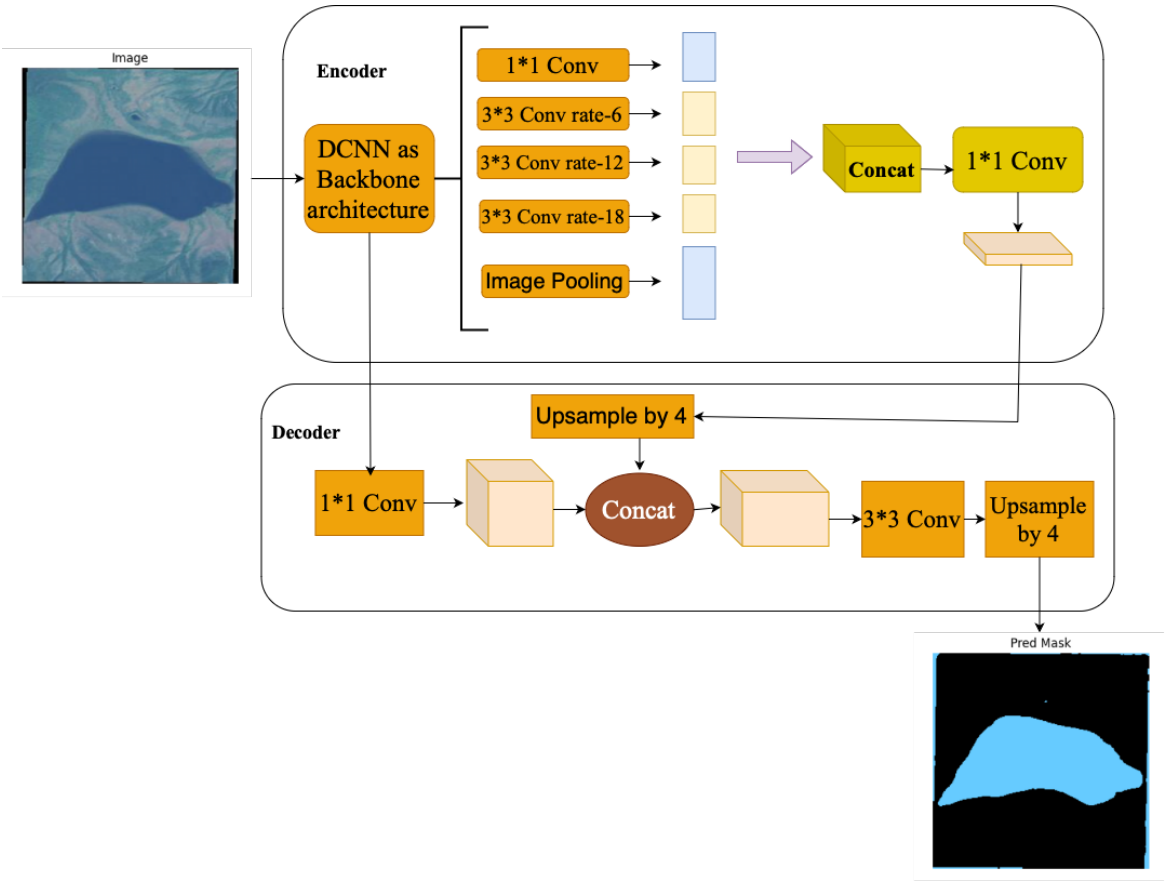


Figure 3.3: Proposed DeepLabv3+ Model with DCNN as Backbone

3.2.2.1 Key Components of the Deeplabv3+ Model

1. **Atrous (Dilated) Convolutions:** DeepLabV3+ with a DCNN backbone utilizes atrous convolutions to capture multi-scale contextual information without down-sampling the feature maps.Atrous convolutions with different dilation rates are employed to capture features at multiple scales, enabling the model to effectively segment objects of various sizes.

2. **Atrous Spatial Pyramid Pooling (ASPP):** The ASPP module aggregates contextual information from multiple scales using atrous convolutions with different rates.It enhances the network's ability to capture both local and global context, which is crucial for accurate semantic segmentation.

3. **Decoder Module:** The decoder module refines the segmentation predictions obtained from the ASPP module by restoring spatial details and integrating high-level semantic information.t includes upsampling layers to restore the spatial resolution of the feature maps and fusion layers to combine semantic information from the ASPP module with detailed spatial information.

4. **Upsampling:** It is the Bilinear interpolation to upsample the feature maps.Skip connections from earlier layers in the encoder to combine high-resolution features.This module upsamples the low-resolution feature maps obtained from the encoder to the original image resolution.

5. **Final Classification Layer:** It applies a 1x1 convolution to produce the final segmentation map with the desired water bodies and non-water bodies classes.The final classification layer produces pixel-wise segmentation predictions. It typically consists of a convolutional layer with softmax activation, assigning a probability distribution to each pixel across different classes e.g., water bodies and non-water bodies.The segmentation masks generated by this layer delineate the extent of water regions in the input imagery.

By integrating these key components, DeepLabV3+ with a ResNet-50 backbone offers a powerful framework for semantic segmentation, enabling accurate identification and delineation of flood water areas in satellite or aerial imagery.

3.2.3 ResNet50V2 Backbone Architecture

ResNet50 is a widely used convolutional neural network (CNN) that consists of 50 layers. It is part of the ResNet (Residual Network) family [20].The key innovation of ResNet is the introduction of residual learning, which helps in training very deep networks by mitigating the vanishing gradient problem.It is a crucial

network for us to understand. It is the basis of much academic research in this field [21]. The ResNet50V2 model serves as a robust feature extractor in our water bodies segmentation task. By leveraging its deep architecture and pre-trained weights,we can effectively extract rich and hierarchical feature representations from input images, which are crucial for achieving accurate segmentation results in the subsequent stages of our model.

Here is the model architecture of the ResNet50V2 Deep Convolutional neural network in 3.4

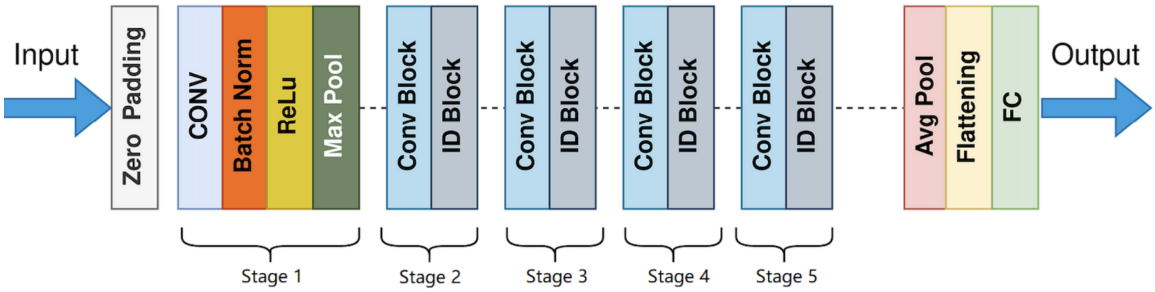


Figure 3.4: ResNet50 Model Architecture

3.2.3.1 Key Components of ResNet50V2

1. **Convolutional Layers:** These layers apply convolution operations to the input images, extracting features through filters that learn spatial hierarchies of features from low to high levels.

2. **Residual Blocks:** Residual blocks are the fundamental building units of ResNet. They contain skip connections (or shortcut connections) that bypass one or more layers, allowing the gradient to flow directly through the network, thus facilitating the training of deeper networks.

3. **Batch Normalization:** Batch normalization is applied after each convolution layer and before the activation function to stabilize and accelerate the training process by normalizing the inputs to each layer.

4. **Activation Functions:** ReLU (Rectified Linear Unit) is used as the activation function, introducing non-linearity into the model and helping it learn complex patterns.

5. **Pooling Layers:** Max pooling and average pooling layers are used to downsample the spatial dimensions of the feature maps, reducing the computational complexity and helping in achieving spatial invariance.

6. **Fully Connected Layers:** In the context of using ResNet-50 as a feature extractor, the fully connected layers (used for classification in the original network) are typically discarded, and the focus is on using the convolutional layers to extract meaningful feature representations.

3.2.3.2 Implementation of ResNet50 as a Feature Extractor

When used as a feature extractor, ResNet50 is responsible for learning and providing rich feature representations from the input images.

1. **Loading the Pre-trained Model:** The pre-trained ResNet-50 model is loaded with weights trained on the ImageNet dataset. This leverages transfer learning, where the model's ability to recognize general features from a large and diverse dataset (ImageNet) is utilized for our specific task of flood water segmentation.

2. **Extracting Features:** The feature maps produced by the ResNet-50 backbone are used as input to the DeepLabV3+ architecture. The include-top=False argument ensures that the fully connected layers are not included, focusing on convolutional layers for feature extraction.

3. **Integration with DeepLabV3+:** The extracted features are fed into the DeepLabV3+ model, which performs further processing, including atrous spatial pyramid pooling and upsampling, to produce the final segmentation map.

3.2.4 MobileNetV3 Backbone Architecture

MobileNetV3 is a highly efficient neural network architecture designed for mobile applications, making it an excellent choice as a backbone for advanced segmentation models like DeepLabV3+. Its design principles and innovations ensure that it delivers strong performance while being computationally feasible for embedded devices. By combining MobileNetV3 with DeepLabV3+, we benefit from both

the efficiency and powerful feature extraction capabilities of MobileNetV3 and the advanced segmentation techniques of DeepLabV3+, making it well-suited for tasks like water body segmentation.

3.2.4.1 Key Components of MobileNetV3

1. **Depthwise Separable Convolutions:** These convolutions decompose standard convolutions into two simpler operations: depthwise convolutions (which apply a single filter per input channel) and pointwise convolutions. This reduces the number of parameters and computational complexity significantly, making the network faster and lighter.

2. **Inverted Residuals with Linear Bottlenecks:** This structure preserves information during the transformation, leading to better gradient flow and higher efficiency.

3. **Squeeze-and-Excitation (SE) Blocks:** SE blocks help the network to focus on the most informative features, improving the network's representational power.

4. **Neural Architecture Search (NAS):** NAS ensures that the architecture is tailored for both efficiency and high performance, making it suitable for deployment on resource-constrained devices.

3.2.4.2 Implementation of MobileNetV3 as Feature Extractor

Using MobileNetV3 as a feature extractor involves leveraging the pre-trained network's ability to generate rich, hierarchical feature representations from images. This process is essential in transfer learning, where a pre-trained model is adapted to a new task.

1. **Semantic Segmentation:** In semantic segmentation, the feature maps extracted from MobileNetV3 can be used as input to a segmentation head (e.g., ASPP in DeepLabV3+). The segmentation head processes these features to generate a pixel-wise classification map of the input image.

2. **Efficiency:** MobileNetV3 is designed to be lightweight and computationally efficient, making it suitable for mobile and embedded devices.

3. **Transfer Learning:** Leveraging pre-trained weights from MobileNetV3 allows for effective transfer learning. This means you can benefit from the knowledge the network has already learned from a large dataset like ImageNet.

3.2.5 EfficientNetV2 Backbone Architecture

EfficientNetV2 is a family of convolutional neural networks designed to provide state-of-the-art performance while being computationally efficient.

3.2.5.1 Key Components of EfficientNetV2

1. **Fused-MBConv Blocks:** The Fused-MBConv block is a new building block in EfficientNetV2, which combines the strengths of MobileNetV2's MBConv blocks and traditional convolutions.

   MBConv is an inverted residual block with a depthwise separable convolution and an expansion phase. Fused Convolution Combines expansion and depthwise convolution into a single step to reduce memory access and improve efficiency.

   It reduces the number of operations and memory footprint, making the model faster and more efficient.

2. **Compound Scaling:** Achieves a better trade-off between model size, speed, and accuracy compared to scaling only one or two dimensions.

3. **Multi-Scale Features:** EfficientNetV2 is designed to capture multi-scale features efficiently, which is crucial for tasks like object detection and semantic segmentation.

3.2.5.2 Implementation of EfficientNetV2 as Feature Extractor

Using EfficientNetV2 as a feature extractor involves leveraging its pre-trained convolutional layers to generate rich feature representations from input images.

1. **Load the Pre-trained EfficientNetV2 Model:** EfficientNetV2 can be loaded with pre-trained weights, usually on ImageNet to take advantage of the learned features.

2. **Extract Features:** The output from the last convolutional layer (or an intermediate layer) of EfficientNetV2 is used as the feature representation of the input image.

3.2.6 Training Details

1. **Hyperparameters and Configurations:**

   - **Batch Size:** The number of samples processed before the model is updated. A batch size of 8 was chosen to balance memory usage and convergence speed.

   - **Epochs:** The number of complete passes through the training dataset. The model was trained for 45 epochs to ensure sufficient learning.

   - **Learning Rate:** The step size at each iteration while moving toward a minimum of the loss function. An initial learning rate of 0.001 was used, with a learning rate scheduler to reduce the rate.

2. **Optimizer**

   - **Adam Optimizer:** Combines the advantages of two other extensions of stochastic gradient descent. Adam optimizer was chosen for its efficiency and effectiveness in handling sparse gradients.

3. **Callbacks**

   - **Early Stopping:** Monitors the validation loss and stops training when it has not improved for 10 consecutive epochs. This helps prevent overfitting and unnecessary training time.

   - **Model Checkpoint:** Saves the best model based on validation accuracy. This ensures the best performing model is retained.

   - **Learning Rate Scheduler:** Reduces the learning rate by a factor

of 0.1 if the validation loss does not improve for 5 consecutive epochs. This helps in fine-tuning the model during training.

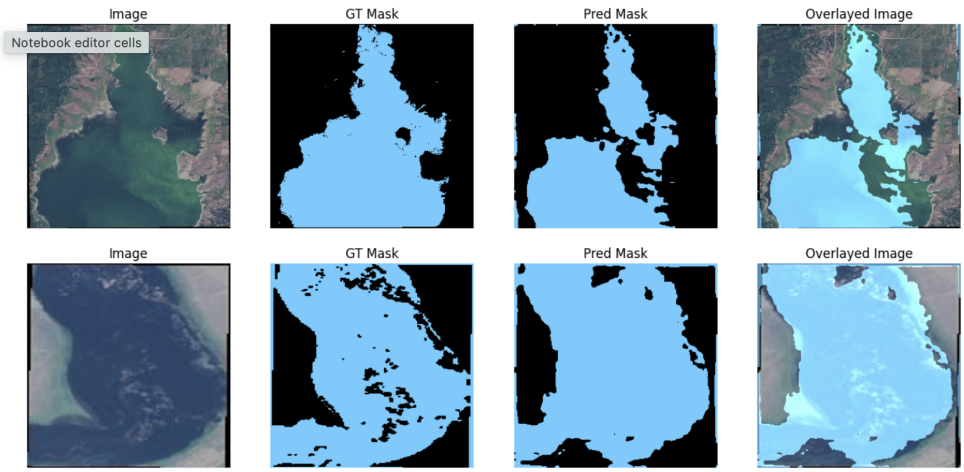
4. **Normalization and Regularization**

   - **Batch Normalization:** Applied to improve the training speed and stability of the neural network. It normalizes the input of each layer to have zero mean and unit variance.

3.2.7 Evaluation of the Model

- The training was monitored using the validation dataset to adjust parameters and prevent overfitting.

These configurations and techniques collectively ensure the model's effective training and generalization for the task of water bodies segmentation. The figure 3.5 below shows the images after training and evaluation of the remote sensing images.



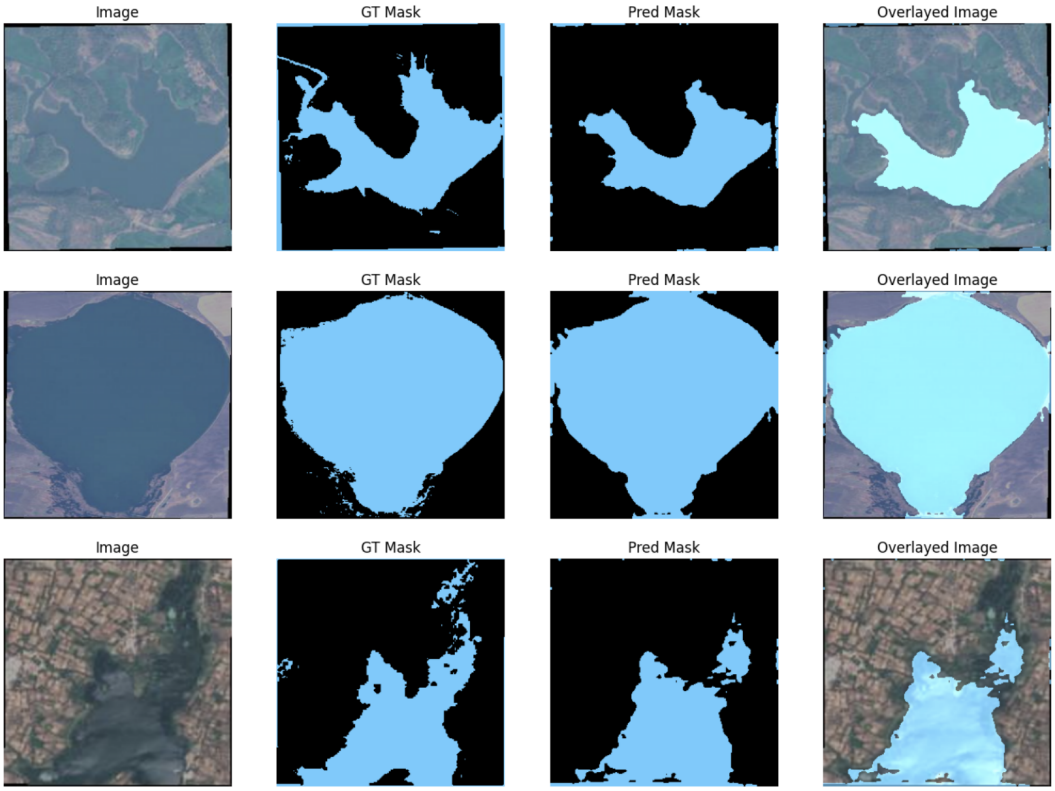


Figure 3.5: Result of the input images after evaluation

3.3 Conclusion

In this chapter, a detailed explanation of the proposed methodology is provided and then demonstrated its implementation. The full methodology was divided into four parts: Data preprocessing,Deeplabv3+ model with DCNNs as backbone, Training Details and Evaluation the model. The following chapter is organised with the results and outcomes of our proposed model with necessary explanation and analysis. The results and effectiveness of our method will be described in the following chapter.

Chapter 4

Results and Discussions

4.1 Introduction

We will talk about the data set utilized for the thesis work and its performance for various models in this part. This chapter will also cover the end results and their analysis. The primary objective of this research is to develop and evaluate an efficient DeepLabV3+ model with different DCNNs as backbone and ensemble them by weighted average method for the task of water bodies segmentation. The dataset used in this study annotated with ground truth masks indicating the presence of water. The research involves preprocessing the data, training the model, and fine-tuning its parameters to achieve optimal performance. Key performance metrics, such as accuracy, precision, recall, F1 score and intersection over union (IoU) are used to evaluate the model's effectiveness.The significance of this thesis is finally examined, and the assessment of this research project is concluded.

4.2 Dataset Description

For our thesis work, we have collected the remote sensing images and their corresponding masks from kaggle [22] [23] which includes a total of 2900 images.The datasets contains of two classes,which are water bodies areas, and non-water bodies areas.The dataset utilized in this thesis comprises aerial and satellite images. It is curated to provide a diverse representation of various geographic locations, ensuring the robustness and generalizability of the model. The dataset includes both synthetic and real-world images, with annotations indicating the presence and extent of water.Images are resized to a uniform dimension (e.g. 256x256

pixels) to ensure compatibility with the input requirements of the DeepLabV3+ model.Techniques such as rotation, flipping, and scaling are applied to increase the diversity of the training set and improve the model's generalizability.we divided the dataset into 80% of dataset used to train the model, Comprises 10% of the total dataset, used for hyperparameter tuning and model validation during training and Comprises 10% of the total dataset, used for evaluating the final model's performance on unseen data. After that we increased the dataset using different augmentation techniques.

4.3 Experimental Setup

We trained our all model in Kaggle that provided 16GB RAM and two 18GB T4 GPUs. We resized our image into 256 x 256 pixels image for training our all model and used data augmentation for better performance.

4.4 Evaluation of Performance

4.4.1 Evaluation Metrics

The evaluation metrics are crucial for assessing the performance and accuracy for our model.In order to evaluate our models performance, many assessment measures are employed. These measurements show how one model performs better than another. For our model evaluation, we settled on some metrics: Accuracy, Intersection over Union, Precision, Recall, Loss-function,F1 Score. Four words are presented from the confusion matrix concept: True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP). Below is a description of them :

- **True Positive:** Positive labels that are correctly classified.
- **True Negative:** Negative labels that are correctly classified.
- **False Positive:** Actual negative labels that are incorrectly predicted to be positive.

- **False Negative:** Actual positive labels incorrectly predicted to be negative. Our evaluation metrics are described below by equations based on TP,TN,FP and FN.

Below are the description for all the evaluation metrics that we have used for the measurement of our proposed model.

- **Accuracy:** Accuracy is a common evaluation metric that measures the proportion of correctly predicted instances out of the total instances in a dataset. It provides an overall measure of how well the model predicts the correct class labels.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

- **Precision:** Precision is a metric that quantifies the proportion of correctly predicted positive instances out of the total instances predicted as positive. It focuses on the correctness of positive predictions, indicating how well the model avoids false positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4.2}$$

- **Recall:** Recall, also known as sensitivity or true positive rate, measures the proportion of correctly predicted positive instances out of the actual positive instances in the dataset. It highlights the model's ability to identify true positives and avoid false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4.3}$$

- **F1-Score:** The F1 score is a statistical measure used to evaluate the performance of a classification model. It is the harmonic mean of precision and recall, providing a single metric that balances both the model's accuracy in identifying true positives and its ability to minimize false positives and false negatives.It ranges from 0 to 1, where 1 indicates perfect precision and recall, and 0 indicates the worst performance. It is particularly useful in

situations where you need to balance the trade-off between precision and recall, rather than optimizing for just one.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.4}$$

- **Loss Function:** A loss function calculates the discrepancy between predicted and actual values during model training. It quantifies the error made by the model and guides the optimization process to minimize this error. Different loss functions are used depending on the task, such as mean squared error for regression or cross-entropy for classification.

- **Intersection over Union:** Intersection over Union (IoU) is an evaluation metric used to measure the accuracy of an object detector on a particular dataset. It is commonly used in segmentation tasks to compare the similarity between the predicted segmentation mask and the ground truth mask.

$$\text{IoU} = \frac{TP}{TP + FP + FN} \tag{4.5}$$

The following figure 4.6 shows the overview of the Intersection over Union concept,which is very important for the segmentation task.

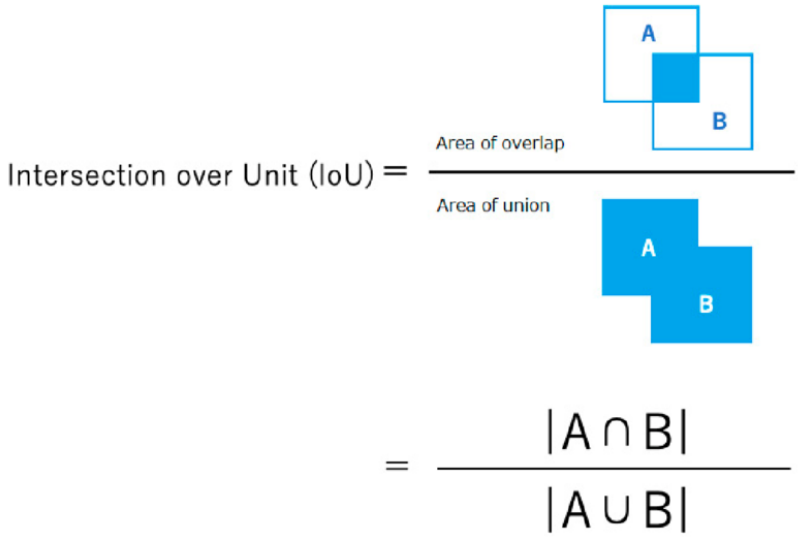


Figure 4.1: Graphical representation of IoU

Where:

- A is the set of predicted positive pixels (the predicted segmentation mask).
- B is the set of actual positive pixels (the ground truth segmentation mask).
- AB is the number of pixels where both the predicted and ground truth masks are positive (the intersection).
- AB is the number of pixels where either the predicted or ground truth masks are positive (the union).

IoU ranges from 0 to 1, where:

- An IoU of 1 indicates perfect overlap between the predicted and ground truth masks.
- An IoU of 0 indicates no overlap between the predicted and ground truth masks.

4.4.2 Comparison of different pre-trained model

Comparing different pre-trained models for water bodies segmentation can provide valuable insights into their performance, robustness, and suitability for our specific application. The different pre-trained DCNNs were tried such as VGG-19, ResNet50V2, DenseNet, MobileNetV2, EfficientNetV2 etc. The results are demonstrated in the following table:

Table 4.1: Different Pre-trained Models as Backbone for Water bodies Segmentation

Model	**Backbone**	**Encoder**	**Decoder**	**Performance**
DeepLabV3+	ResNet50V2	Atrous convolutions	ASPP, skip connections	High
DeepLabV3+	MobileNetV3	Depthwise separable convolutions	ASPP, skip connections	High
DeepLabV3+	VGG-19	Convolutional layers	ASPP, skip connections	Low
DeepLabV3+	DenseNet	Convolutional layers	ASPP, skip connections	Low
DeepLabV3+	EfficientNetV2	Convolutional layers	ASPP, skip connections	Moderate

From the above table, we can clearly see that the ResNet50V2 and MobileNetV3 gives better results than other pre-trained models. Now the actual results of different pre-trained models and our proposed average weighted ensemble model are in the following table:

Table 4.2: Results by using Different Pre-trained Models as Backbone

Model Name	**Accuracy**	**Precision**	**Recall**	**IoU**
VGG-19	71%	47%	-	57%
DenseNet	75%	34%	-	54.7%
MobileNetV3	92.6%	87.8%	87%	78%
EfficientNetV2	92%	91%	85%	79%
ResNet50V2	91.3%	87%	85.7%	76.6%
Proposed Ensemble Model	**96%**	**94%**	**93%**	**88%**

4.4.3 Result Analysis of the Ensemble model

The ensemble model for water bodies segmentation, using DeepLabV3+ with MobileNetV3, EfficientNetV2, and ResNet50V2 as backbones, significantly improves segmentation accuracy and robustness. The combination of these models through weighted averaging leverages their individual strengths, resulting in superior performance compared to single models. This approach demonstrates the effectiveness of ensemble learning in enhancing the accuracy and reliability of semantic segmentation tasks. Utilizing a dataset comprising images of water bodies alongside corresponding masks delineating water regions, we meticulously preprocessed the data by resizing images and masks to a standard size. As we have observed, ResNet50V2 performed the best, followed by MobileNetV3 and then EfficientNetV2. Consequently, we assigned greater weight to the DeepLabV3+ model using ResNet50V2 as the backbone during the training period. Specifically, we used a weight of 0.5 for the ResNet50V2 backbone, 0.3 for the MobileNetV3 backbone, and 0.2 for the EfficientNetV2 backbone.

Our proposed ensemble model achieved an accuracy of 94% . Besides we have achieved higher precision of 92% and recall is 90% .

Figure 4.2, 4.3, 4.4 , 4.5 and 4.6 represents the Accuracy, loss curve, precision, F-1 Score and IoU etc respectively of our proposed ensemble model.

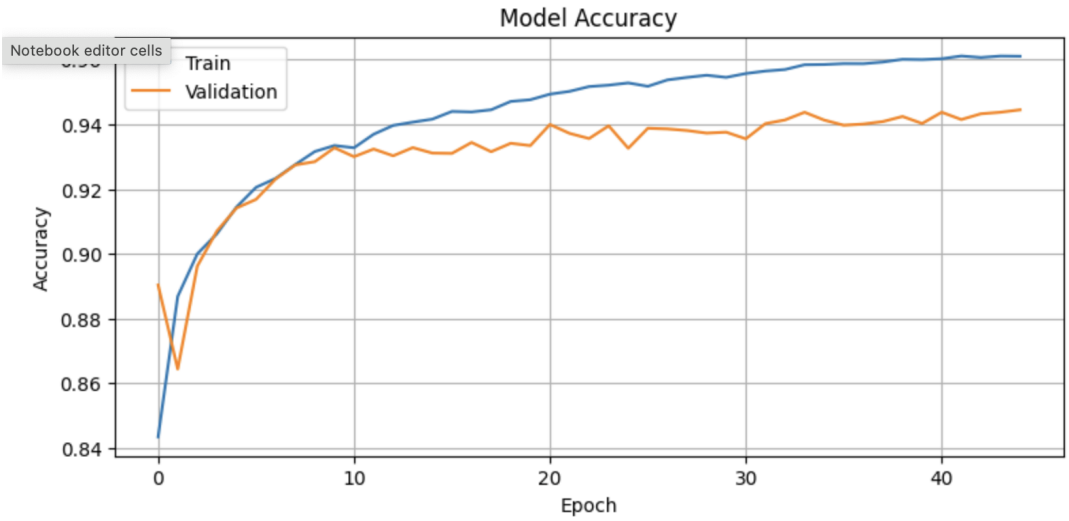


Figure 4.2: Accuracy Curve

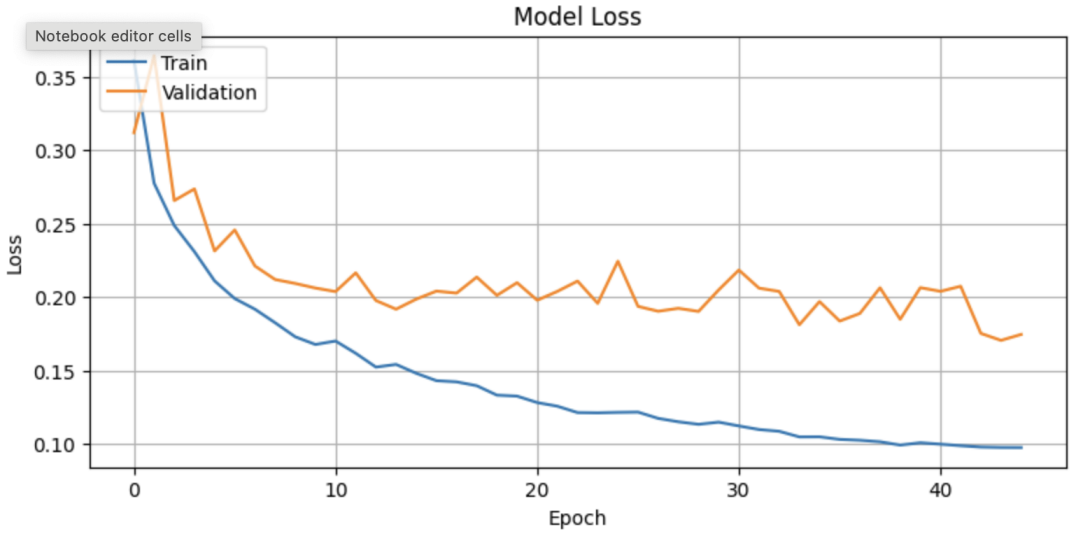


Figure 4.3: Loss Curve

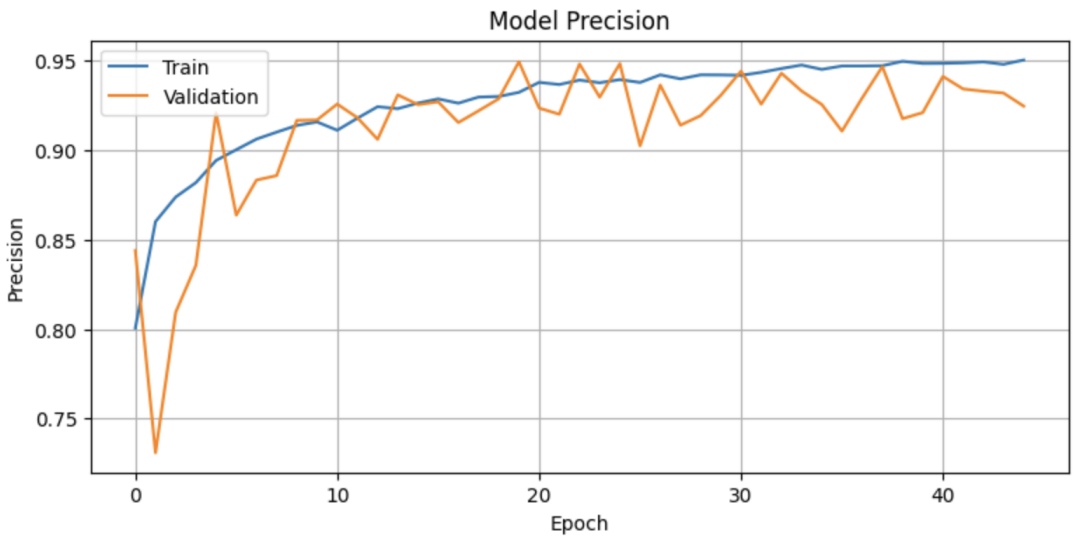


Figure 4.4: Precision Curve

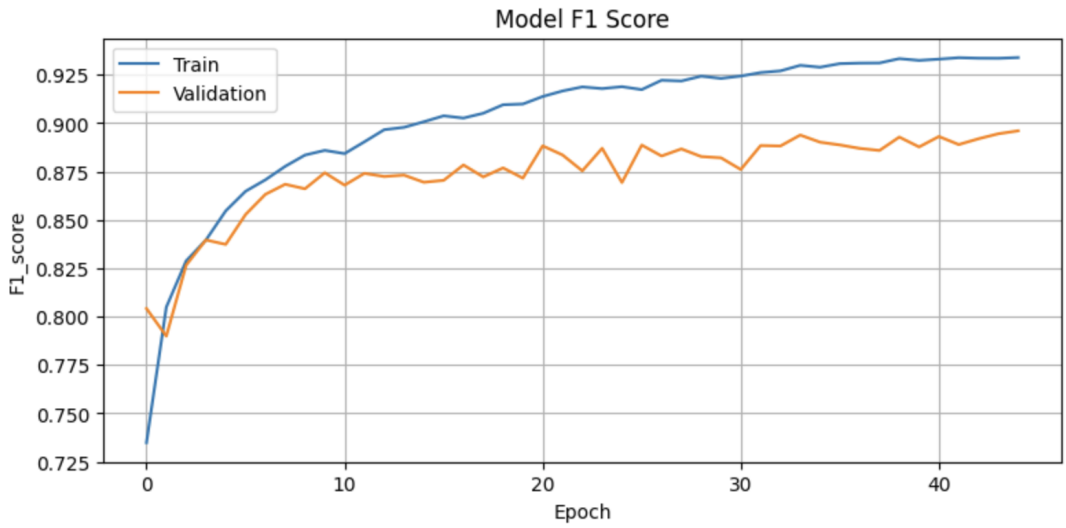


Figure 4.5: F-1 Score Curve

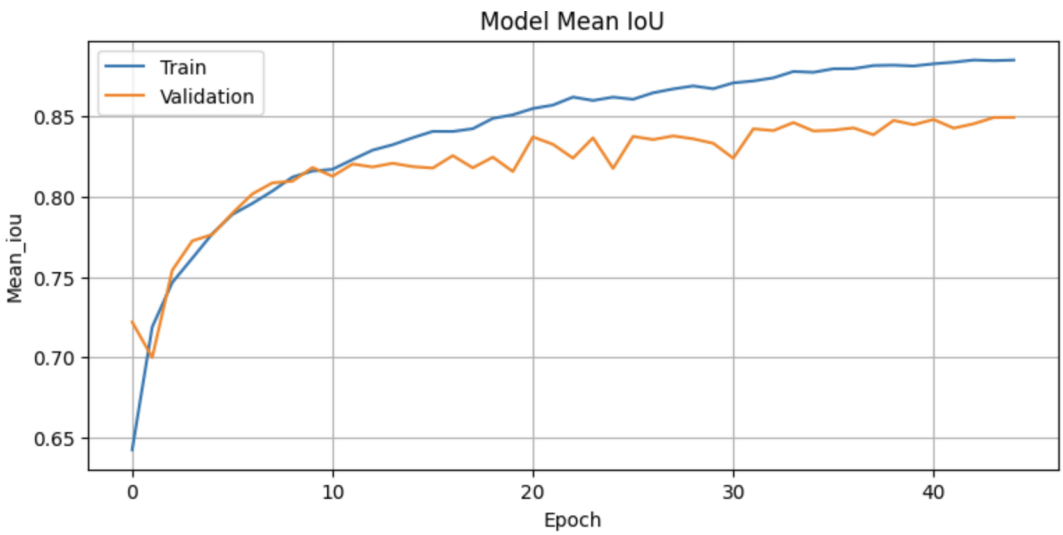


Figure 4.6: mIoU Curve

4.4.4 Performance Evaluation in Relation to Previous Literature

Different research work was mentioned in the literature review. Table 4.8 shows the different scenario among all of the works with our proposed methods.From the table we can clearly see that our model performed better in the segmentation scenario.

Table 4.3: Comparison among different existing methods and proposed method

Methods	**Accuracy**	**Precision**	**Recall**
H.Rizk et al.[24]	73.42%	73%	75%
DP Singh et al.[25]	76.62%	93.49%	79.44%
F.Pech-may et al.[26]	81%	84%	78%
A.jamali et al.[27]	95%	92%	69.7%
T.Ahmed [28]	80.7%	82%	79%
Proposed method	**96%**	**94%**	**93%**

4.5 Impact Analysis

The research presented in our thesis on using a weighted average ensemble model for water bodies segmentation holds significant impact across various dimensions, including scientific, practical, and environmental aspects. From a scientific perspective, our work advances the field of semantic segmentation by demonstrating how an ensemble approach can enhance model accuracy and robustness. By integrating multiple backbones (ResNet50V2, MobileNetV3, EfficientNetV2) within the DeepLabV3+ framework, our ensemble model achieves higher segmentation accuracy compared to individual models. This advancement contributes to the existing body of knowledge on deep learning techniques, specifically highlighting the effectiveness of ensemble methods in improving performance. Additionally, the use of a weighted average method for combining predictions, based on the performance of each backbone, represents a methodological innovation that can inspire further research and application in various segmentation tasks. Practically, our thesis provides valuable improvements in the domain of water body monitoring. The enhanced accuracy of our ensemble model facilitates better detection and mapping of water bodies, which is crucial for applications in hydrology, environmental monitoring, and land management. Accurate segmentation of water bodies supports effective water resource management by enabling precise monitoring of water levels, distribution, and changes over time. This can inform decision-making processes related to water conservation, flood management, and ecological preservation. Environmentally, the accurate segmentation of water bodies has a direct impact on ecological studies and conservation efforts. By providing reliable data on the extent and condition of water bodies, our research aids in assessing the health of aquatic ecosystems, tracking the effects of

climate change, and implementing sustainable environmental practices. Overall, our thesis not only advances the technical aspects of semantic segmentation but also contributes significantly to practical and environmental applications, highlighting the broad relevance and importance of our research.

4.5.1 Social Environmental Impact

Our thesis on using a weighted average ensemble model for water bodies segmentation carries significant social and environmental implications, influencing various aspects of society and the environment.Here are a few key aspects:

1. **Environmental Conservation:** The accurate segmentation of water bodies enabled by our ensemble model contributes directly to environmental conservation efforts. By providing precise data on the extent and condition of water bodies, our research supports the assessment and monitoring of aquatic ecosystems. This information is crucial for understanding ecological health, identifying areas of concern, and implementing targeted conservation strategies.

2. **Climate Change Resilience:** In the face of climate change, reliable monitoring and management of water bodies are essential. Our thesis enhances the capability to monitor changes in water levels, distribution patterns, and ecosystem dynamics over time. This data is invaluable for adapting to climate variability, predicting and managing flood risks, and safeguarding communities vulnerable to water-related disasters.

3. **Resource Management and Sustainability:** Effective water resource management is critical for ensuring sustainable development. Our ensemble model improves the accuracy of water body segmentation, thereby supporting informed decision-making in water allocation, agricultural practices, and industrial operations. This contributes to optimizing resource use, minimizing water wastage, and promoting sustainable development practices that balance environmental conservation with socio-economic needs.

4. **Public Awareness and Education:** By demonstrating the application of advanced segmentation techniques in environmental monitoring, our thesis

raises public awareness about the importance of water conservation and ecosystem health. It underscores the role of technology in addressing environmental challenges and encourages stakeholders to actively participate in conservation efforts.

5. **Policy and Governance:** The insights and data generated from our research can inform policy development and governance frameworks related to water management and environmental protection. Our findings provide evidence-based recommendations for policy-makers, empowering them to enact legislation and regulations that prioritize environmental sustainability, protect water resources, and promote ecosystem resilience.

4.5.2 Economic impact

The economic impact of this research on water bodies segmentation is substantial. Here are a few key points:

1. **Optimized Resource Allocation:** Accurate segmentation of water bodies facilitated by our ensemble model enhances efficiency in resource allocation. Industries reliant on water resources, such as agriculture, manufacturing, and energy production, can benefit from precise data on water availability and distribution. This enables them to optimize water use, reduce operational costs associated with water management, and mitigate risks related to water scarcity or overuse.

2. **Infrastructure Planning and Investment:** Urban and rural development planning relies heavily on accurate environmental data, including information on water bodies. Our thesis provides reliable insights into the spatial distribution and characteristics of water bodies, aiding in infrastructure planning and investment decisions. Governments and private developers can use this information to strategically allocate resources for water infrastructure projects, such as reservoirs, dams, and drainage systems.

3. **Environmental Impact Assessment (EIA) and Compliance:** n many

regions, regulatory frameworks require rigorous environmental impact assessments (EIAs) for development projects affecting water bodies. Our ensemble model improves the accuracy and reliability of environmental data used in EIAs, ensuring compliance with regulatory standards and minimizing potential environmental liabilities.

4. **Eco-Tourism and Recreation:** Accurate segmentation and monitoring of water bodies support eco-tourism and recreational activities that contribute to local economies. Tourist destinations centered around lakes, rivers, and coastal areas benefit from reliable information on water quality and accessibility. Our research enhances the attractiveness of these destinations by promoting sustainable management practices and preserving natural landscapes.

5. **Innovation and Technological Advancement:** The application of advanced segmentation techniques in environmental monitoring demonstrates technological innovation and fosters growth in the technology sector. Our thesis showcases the capabilities of deep learning models in addressing complex environmental challenges, stimulating research and development in artificial intelligence (AI) and machine learning (ML) applications for environmental sciences.

4.5.3 Ethical Impact

The ethical impact of this research is significant and raises important considerations. Here are a few key points:

1. **Environmental Justice and Equity:** Accurate segmentation of water bodies through our ensemble model contributes to environmental justice by ensuring fair access to clean and sustainable water resources for all communities. By providing precise data on water distribution and quality, our research supports efforts to address disparities in water access and mitigate environmental burdens disproportionately borne by marginalized or vulnerable populations.

2. **Ecosystem Preservation and Conservation Ethics:** The segmentation and monitoring of water bodies facilitated by our research uphold principles of ecosystem preservation and conservation ethics. By accurately delineating water bodies and monitoring their health, our model aids in safeguarding biodiversity, protecting fragile ecosystems, and preserving natural habitats.

3. **Ethical Use of Technology and Data Privacy:** As we leverage advanced segmentation techniques and AI-driven models, ethical considerations regarding technology use and data privacy become paramount. Our thesis emphasizes ethical practices in data collection, processing, and storage to protect individual privacy rights and prevent misuse of sensitive environmental data.

4. **Community Engagement and Empowerment:** Engaging communities in environmental monitoring and decision-making processes is essential to ethical practice. Our research encourages community empowerment through participatory approaches, involving stakeholders in data collection, interpretation, and implementation of conservation measures. By fostering collaboration and inclusivity, our thesis promotes ethical principles of environmental governance, where communities have a voice in shaping policies and practices that affect their environment and well-being.

5. **Transparency and Accountability in Environmental Management:** Our thesis promotes transparency and accountability in environmental management practices by providing stakeholders with reliable information on water resources. Transparent data-driven decisions, facilitated by accurate segmentation, enhance public trust in regulatory processes, project planning, and resource allocation.

4.6 Conclusion

In this chapter, we dive into the details of the data set used in our work. We then proceed to emphasize the impact our thesis has on various aspects of society, including the social, environmental, and ethical spheres. Finally, we provide a thorough evaluation of the entire thesis.The model showcased commendable performance across various evaluation metrics, indicating its effectiveness in accurately identifying water areas. With an overall accuracy of 96%, the model demonstrated robustness in delineating water regions. Precision, Recall, and F1-score metrics further corroborated the model's accuracy and reliability in water bodies segmentation tasks. Additionally, achieving an overall IoU of 88% underscored the model's ability to accurately align predicted masks with ground truth water areas. The fine-tuning process yielded notable improvements in performance metrics, affirming the efficacy of refining the model's parameters to better suit the task at hand.

Chapter 5

Conclusion

5.1 Conclusion

This thesis aimed to address the challenge of accurately segmenting water bodies in very high-resolution satellite and aerial images, a task of paramount importance for applications such as environmental monitoring, urban planning, and disaster management. Traditional segmentation methods have often struggled with the complexities and variability of such high-resolution data. To overcome these challenges, we leveraged the advanced DeepLabv3+ model, enhanced with an ensemble of state-of-the-art feature extraction backbones: ResNet50V2, MobileNetV3, and EfficientNetV2. By employing a weighted averaging ensemble technique, we successfully integrated the strengths of these three backbones within the encoder part of the DeepLabv3+ model. This ensemble approach significantly improved the feature extraction process, leading to enhanced segmentation performance. Our experimental results demonstrated that the proposed ensemble model achieved approximately 96% accuracy and a mean Intersection over Union (mean IoU) of 88%, along with notable improvements in the F1-score, precision, and recall metrics. Extensive experiments validated the effectiveness of the ensemble approach, showcasing significant improvements across multiple evaluation metrics. The improved segmentation model provides a more accurate and efficient tool for environmental monitoring and management. Despite these promising results, this research also encountered certain challenges and limitations. The complexity of the ensemble model increased computational requirements, and further optimization could be explored to enhance efficiency. Additionally, the generalizability of the model to different types of remote sensing images and varying environmental conditions could be investigated in future work.

5.2 Future Work

While the proposed methodology for water bodies segmentation using DeepLabV3+ with a weighted averaging ensemble method has shown promising results, there are several avenues for future research and improvement:

1. **Model Optimization:** Future work could focus on optimizing the ensemble model to reduce its computational complexity and improve efficiency. Techniques such as model pruning, quantization, and knowledge distillation could be employed to decrease the model size and increase inference speed without sacrificing accuracy.

2. **Real-Time Implementation :** One of the key future directions is optimizing the model for real-time applications. This involves reducing the computational complexity and improving the processing speed without significantly compromising the accuracy of segmentation.

3. **Semi-Supervised and Unsupervised Learning :** Given the challenges associated with obtaining large amounts of annotated water bodies data, exploring semi-supervised and unsupervised learning techniques could be beneficial. These methods can leverage unlabelled data to improve model performance and robustness.

4. **Expansion of Dataset :** Expanding the dataset to include a wider variety of water bodies scenarios, geographical locations, and environmental conditions will enhance the model's generalizability. Incorporating data from different seasons and weather conditions can further improve the model's ability to handle diverse real-world situations.

5. **Integration with Other Data Sources :** Combining satellite imagery with other data sources such as social media feeds, weather forecasts, and IoT sensors could provide a more comprehensive prediction system. Multimodel data fusion techniques can be investigated to integrate these diverse data sources effectively.

6. **Evaluation of Model Robustness :** Conducting extensive evaluations

to assess the robustness of the model under various conditions, including extreme weather events and low-quality satelite imagery, can help identify potential weaknesses and areas for improvement.

7. **Collaborative Research :** Engaging in collaborative research with governmental agencies, non-profit organizations, and other academic institutions can provide access to additional resources, data, and expertise. This collaboration can facilitate the development of more comprehensive and impactful solutions.

References

[1] M. Zaffaroni, C. Rossi *et al.*, 'Water segmentation with deep learning models for flood detection and monitoring,' *International Journal of Information Systems for Crisis Response and Management*, pp. 24–27, 2020 (cit. on p. 1).

[2] M. Wieland, S. Martinis, R. Kiefl and V. Gstaiger, 'Semantic segmentation of water bodies in very high-resolution satellite and aerial images,' *Remote Sensing of Environment*, vol. 287, p. 113 452, 2023 (cit. on p. 1).

[3] B. Bahrami and H. Arbabkhah, 'Enhanced flood detection through precise water segmentation using advanced deep learning models,' *J. Civ. Eng. Res*, vol. 6, pp. 1–8, 2024 (cit. on p. 1).

[4] N. A. Muhadi, A. F. Abdullah, S. K. Bejo, M. R. Mahadi and A. Mijic, 'Deep learning semantic segmentation for water level estimation using surveillance camera,' *Applied Sciences*, vol. 11, no. 20, p. 9691, 2021 (cit. on p. 1).

[5] J. Kim, H. Jeon and D.-j. Kim, 'Extracting flooded areas in southeast asia using segnet and u-net,' *Korean Journal of Remote Sensing*, vol. 36, no. 5_3, pp. 1095–1107, 2020 (cit. on p. 2).

[6] S. Kadry, M. A. Al-Betar, S. Yassine, R. Mohan, R. Arunmozhi and V. Rajinikanth, 'Automatic detection of waterbodies from satellite images using deeplabv3+,' in *International Conference on Mining Intelligence and Knowledge Exploration*, Springer, 2023, pp. 78–86 (cit. on p. 2).

[7] T. Akiyama *et al.*, 'Deep learning applied to water segmentation,' *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 1189–1193, 2020 (cit. on p. 2).

[8] S. Lv, L. Meng, D. Edwing, S. Xue, X. Geng and X.-H. Yan, 'High-performance segmentation for flood mapping of hisea-1 sar remote sensing images,' *Remote Sensing*, vol. 14, no. 21, p. 5504, 2022 (cit. on pp. 2, 10).

[9] C. Sazara, M. Cetin and K. M. Iftekharuddin, 'Detecting floodwater on roadways from image data with handcrafted features and deep transfer learning,' in *2019 IEEE intelligent transportation systems conference (ITSC)*, IEEE, 2019, pp. 804–809 (cit. on p. 11).

[10] S.-W. Lo, J.-H. Wu, F.-P. Lin and C.-H. Hsu, 'Visual sensing for urban flood monitoring,' *Sensors*, vol. 15, no. 8, pp. 20 006–20 029, 2015 (cit. on p. 11).

[11] A. Rankin and L. Matthies, 'Daytime water detection based on color variation,' in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2010, pp. 215–221 (cit. on p. 11).

[12] N. M. Robertson and T. Chan, 'Aerial image segmentation for flood risk analysis,' in *2009 16th IEEE International Conference on Image Processing (ICIP)*, IEEE, 2009, pp. 597–600 (cit. on p. 11).

[13] M. Mousa, X. Zhang and C. Claudel, 'Flash flood detection in urban cities using ultrasonic and infrared sensors,' *IEEE Sensors Journal*, vol. 16, no. 19, pp. 7204–7216, 2016 (cit. on p. 11).

[14] P. V. K. Borges, J. Mayer and E. Izquierdo, 'A probabilistic model for flood detection in video sequences,' in *2008 15th IEEE International Conference on Image Processing*, IEEE, 2008, pp. 13–16 (cit. on p. 12).

[15] S. Sarp, M. Kuzlu, M. Cetin, C. Sazara and O. Guler, 'Detecting floodwater on roadways from image data using mask-r-cnn,' in *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, 2020, pp. 1–6 (cit. on p. 12).

[16] P. Santana, R. Mendonça and J. Barata, 'Water detection with segmentation guided dynamic texture recognition,' in *2012 IEEE international conference on robotics and biomimetics (ROBIO)*, IEEE, 2012, pp. 1836–1841 (cit. on p. 12).

[17] M. Geetha, M. Manoj, A. Sarika, M. Mohan and S. N. Rao, 'Detection and estimation of the extent of flood from crowd sourced images,' in *2017 international conference on communication and signal processing (ICCSP)*, IEEE, 2017, pp. 0603–0608 (cit. on p. 12).

[18] F. Pech-May, R. Aquino-Santos, O. Álvarez-Cárdenas, J. L. Arandia and G. Rios-Toledo, 'Segmentation and visualization of flooded areas through sentinel-1 images and u-net,' *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2024 (cit. on p. 13).

[19] F. Safavi, T. Chowdhury and M. Rahnemoonfar, 'Comparative study between real-time and non-real-time segmentation models on flooding events,' in *2021 IEEE International Conference on Big Data (Big Data)*, IEEE, 2021, pp. 4199–4207 (cit. on p. 13).

[20] K. He, X. Zhang, S. Ren and J. Sun, 'Deep residual learning for image recognition,' in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778 (cit. on p. 19).

[21] B. Koonce and B. Koonce, 'Resnet 50,' *Convolutional neural networks with swift for tensorflow: image recognition and dataset categorization*, pp. 63–72, 2021 (cit. on p. 20).

[22] M. F. Karim, *Flood area segmentation*, Accessed: 2024-05-31, 2022. [Online]. Available: `https://www.kaggle.com/datasets/faizalkarim/flood-area-segmentation` (cit. on p. 27).

[23] F. Escobar, *Satellite images of water bodies*, Accessed: 2024-05-31, 2022. [Online]. Available: `https://www.kaggle.com/datasets/franciscoescobar/satellite-images-of-water-bodies` (cit. on p. 27).

[24] H. Rizk, Y. Nishimur, H. Yamaguchi and T. Higashino, 'Drone-based water level detection in flood disasters,' *International journal of environmental research and public health*, vol. 19, no. 1, p. 237, 2021 (cit. on p. 36).

[25] A. Rambhad, D. P. Singh and J. Choudhary, 'Detection of flood events from satellite images using deep learning,' in *International Conference on Frontiers of Intelligent Computing: Theory and Applications*, Springer, 2022, pp. 259–268 (cit. on p. 36).

[26] F. Pech-May, J. V. Sanchez-Hernández, L. A. López-Gómez, J. Magaña-Govea and E. M. Mil-Chontal, 'Flooded areas detection through sar images and u-net deep learning model,' *Computación y Sistemas*, vol. 27, no. 2, pp. 449–458, 2023 (cit. on p. 36).

[27] A. Jamali, S. K. Roy, L. H. Beni, B. Pradhan, J. Li and P. Ghamisi, 'Residual wave vision u-net for flood mapping using dual polarization sentinel-1 sar imagery,' *International Journal of Applied Earth Observation and Geoinformation*, vol. 127, p. 103 662, 2024 (cit. on p. 36).

[28] T. Ahmed, 'Mapping floodwater from radar imagery,' EasyChair, Tech. Rep., 2022 (cit. on p. 36).