Machine Learning Based Flood Risk Modeling Using Features from Satellite, Socioeconomic, Geographic, and Building Data

by

Anushka Ray

S.B., Computer Science and Engineering, Massachusetts Institute of Technology (2022)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2023

Author .....
Department of Electrical Engineering and Computer Science
January 20, 2023

Certified by .....
John Fernández
Professor
Thesis Supervisor

Accepted by .....
Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Machine Learning Based Flood Risk Modeling Using Features from Satellite, Socioeconomic, Geographic, and Building Data

by

Anushka Ray

Submitted to the Department of Electrical Engineering and Computer Science
on January 20, 2023, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Due to the effects of climate change coupled with increased urbanization, many cities will be experiencing more frequent and intense flooding in the future. As a result, it would be very beneficial for urban planners to have a low-cost and efficient modeling tool that can determine the flood risk at a granular level such as the census tract. Boston is one such coastal urban city that will experience an increase in flooding. Since each census tract in Boston is unique and varies in population and land use, urban planners and policy makers must know which areas in Boston are the most vulnerable to provide them with resources. This research proposes a machine learning based model that evaluates the flood risk of census tracts in Boston. The overall flood risk of a census tract is determined by aggregating relevant features such as land cover data from aerial satellite imagery via semantic segmentation methods, elevation, slope, and flow accumulation. In addition to these flood hazard features, we also integrate flood vulnerability features from socioeconomic data and building information for each census tract.

Thesis Supervisor: John Fernández
Title: Professor

Acknowledgments

I am very grateful to Professor John Fernández of the MIT Environmental Solutions Initiative Lab and the rest of the Urban Metabolism Group for supporting me through my MEng and offering this research opportunity. It was truly a pleasure working with this lab for the past year. This thesis would also not have been possible without Norhan Bayomi, who organized our team meetings, and provided helpful feedback, encouragement, and support throughout my MEng. I would also like to thank my research partner and good friend Kate Xu, who I really enjoyed collaborating with on this research. I am also thankful for all the dedicated professors, mentors, and teachers who have contributed to my growth as a researcher and person, as well as my family who have supported me throughout my academic journey.

Contents

1 Introduction **15**

1.1 Research Overview 15

1.1.1 Problem Statement 16

1.1.2 Research Objectives 16

2 Related Works **21**

2.1 Flood Risk Modeling Approaches 21

2.2 Flood Risk Models that Weight Features 22

2.3 Semantic Segmentation 23

3 Methods **27**

3.1 Overview 27

3.2 Flood Hazard Features 29

3.2.1 Land Cover 29

3.2.2 Elevation 31

3.2.3 Slope 31

3.2.4 Flow Accumulation 32

3.2.5 Weighting Flood Hazard Features 33

3.3 Flood Vulnerability Features 33

3.3.1 Building Features 34

3.3.2 Social Vulnerabilities 36

3.4 Computational Resources 38

3.5 Image Segmentation Datasets 38

3.5.1 ADE20K . . . 38
3.5.2 Drone Dataset . . . 39
3.5.3 Nearmap . . . 40
3.5.4 Apple Maps . . . 41
3.5.5 Combined Dataset . . . 41
3.5.6 Dataset Labeling Process . . . 42
3.6 Image Segmentation Models . . . 43
3.6.1 Segmenter . . . 43
3.6.2 KNet . . . 44
3.6.3 PSPNet . . . 45
3.7 Training Segmentation Models . . . 45
3.7.1 Pre-Processing . . . 46
3.7.2 Experiments and Metrics . . . 46
3.8 Calculating Land Cover . . . 47
3.8.1 Aggregating Image Features Across Census Tracts . . . 47

4 Results **51**
4.1 Semantic Segmentation . . . 51
4.1.1 Performance . . . 52
4.2 Flood Hazard Features . . . 56
4.2.1 Land Cover Results . . . 56
4.2.2 Elevation Results . . . 59
4.2.3 Slope Results . . . 59
4.2.4 Flow Accumulation Results . . . 60
4.2.5 Flood Hazard Weighting . . . 61
4.3 Flood Vulnerability . . . 63
4.3.1 Socioeconomic Data . . . 63
4.3.2 Building Vulnerability . . . 64
4.3.3 Weighting Flood Vulnerability Features . . . 65
4.4 Obtaining Flood Risk . . . 66

4.5 Overall Flood Risk Model Evaluation 68

5 Conclusion **71**

5.1 Discussion 71

5.2 Limitations 72

5.3 Future Work 73

List of Figures

3-1 Diagram displaying the inputs and outputs of the flood risk model . . 29

3-2 How Flow Accumulation is represented in a raster file 32

3-3 Example Images and Segmentations from the ADE20K Dataset . . . 38

3-4 Example Images and Segmentations from the Kaggle Aerial Semantic Segmentation Drone Dataset 39

3-5 Nearmap AI Layer Features as seen on the Nearmap Map Browser. The layers on the left pane are shown in the image, where the color of the building layer is orange, vegetation layer is green, and the surfaces layer is yellow 40

3-6 Example Images and Segmentations from Nearmap 41

3-7 Example Images and Segmentations from Combined Dataset 42

3-8 Segmenter Architecture [28] 44

3-9 KNet Architecture [31] 44

3-10 PSPNet Architecture [33] 45

3-11 Train and Testing Pipeline to Find the Optimal Model for the Land Cover Task 46

3-12 Exporting Georeferenced Images from Nearmap Map Browser. The coordinate projection is in WGS84. The red outlines on the map separate each Boston census tract 48

3-13 Masking Process 50

4-1 IoU Samples [8] 52

4-2 Example Qualitative Results after Evaluating Fine-tuned Models on Test Dataset 54
4-3 Example Segmentations of Boston Aerial Images Produced by Each Finetuned Model 55
4-4 Example Qualitative Results for NYC Boroughs 56
4-5 Boston Census Tracts with Highest Percentage of Each Class 57
4-6 Boston Climate Ready Map Visualizations for each of the top 10 high flood risk (vulnerability + hazard) census tracts [5] 69
4-7 Boston Climate Ready Map Visualizations for each of the top 10 low flood risk (vulnerability + hazard) census tracts [5] 70

List of Tables

1.1 The thesis is organized in the following manner 17

3.1 Possible values that each of the building features can take 34

3.2 Baseline Model Details for Segmenter, KNet and PSPNet 43

4.1 Quantitative Performance of each Baseline model on the Test Dataset 51

4.2 Quantitative Performance of each Finetuned Segmentation Model on the test dataset 53

4.3 Top 5 Boston Census Tracts with Highest Percentage of Each Class . 56

4.4 High and Low Flood Risk Census Tracts Base on Land Cover Results Only 58

4.5 High Risk Census Tracts due to Low Elevation 59

4.6 Top census tracts with highest flood risk due to low slope values . . . 60

4.7 Top census tracts with highest flood risk due to flow accumulation . . 60

4.8 Census Tracts with the Highest and Lowest Susceptibility to Flood Hazard Risk 61

4.9 Census Tracts Most and Least Vulnerable to Flood Risk (Based on Socioeconomic Features Only) 63

4.10 Census Tracts Most and Least Vulnerable to Flood Risk (Based on Building Features Only) 65

4.11 Census Tracts Most and Least Vulnerable to Flood Risk (Based on Combining Building and Socioeconomic Results) 66

4.12 Census Tracts Most and Least Susceptible to Flood Risk (Based on Combining Vulnerability and Hazard Values) 67

Chapter 1

Introduction

1.1 Research Overview

The increased urbanization of the United States has resulted in the removal of soil and vegetation and an increase in concrete spaces in the form of sidewalks and roads. As a result, the volume, frequency, and intensity of flooding in urban areas has increased [18]. The effects of urbanization coupled with rising sea levels and increases in precipitation will result in frequent and severe flooding in the future. The worsening of flood events due to climate change poses threats to vulnerable populations, infrastructure, and transportation. Additionally, seaside urban cities such as Boston are especially susceptible to damages from increased coastal and riverine flooding.

According to a climate vulnerability assessment of Boston, without stormwater system improvements, over 11,000 structures will be adversely affected by stormwater flooding due to rising sea levels and increase in precipitation. Buildings that are connected to a combined sewer system may face wastewater backup problems. Additionally, as a result of pooled water, it may be difficult to access buildings for refuge, and seeking transportation may also be dangerous. The effects of flooding will be especially felt by vulnerable populations such as those who are above the age of 65, children, those with limited English proficiency, those in poverty, people of color, and populations with disabilities [3].

Models that urban planners can use to evaluate the flood risk of specific census

tracts can be very valuable for the preparation of flood events. This thesis proposes a machine learning based flood risk model for Boston census tracts that utilizes features extracted from aerial satellite imagery, socioeconomic datasets, building datasets, and geographical datasets. It is beneficial to develop models that evaluate flood hazard risk at such a granular level because urban planners can focus on high risk areas and build flood mitigation architectures. They can also invest in better resources for those areas such as storm walls, improved drainage systems, elevated residential buildings, increased vegetation and green spaces, and more [18].

1.1.1 Problem Statement

To our knowledge, there is no low-cost tool for policy makers to evaluate flood risk at the census tract level that utilizes both flood hazard features and vulnerability features in the form of socioecnomic and building data. Such a model is important because according to a 2022 study, there are many areas and properties that are outside of the FEMA flood risk zones that may be vulnerable to flooding. Additionally, flood maps that are currently used to determine the flood risk of properties are based on historic flood data. However, due to the rapid effects of climate change, these models may not accurately capture the present and future risk of flooding nor adapt to these effects [13]. Additionally, many flood risk models rely on only geographic and environmental factors, yet a crucial aspect of flood risk includes social vulnerabilities and building data.

1.1.2 Research Objectives

To address the problem, we propose a model that utilizes information from past studies about factors that contribute to an increase in flood risk and aggregates these features to create a tool for policy makers to evaluate flood risk at the census tract level. The key topics of this research are organized in the thesis according to Table 1.1. Please refer to the locations in the table for more detailed explanations of the research objectives.

Table 1.1: The thesis is organized in the following manner

Objective	Location in Thesis
Discussion of Flood Hazard Features (Land Cover, Slope, Elevation, Flow Accumulation)	Section 3.2
Discussion of Flood Vulnerability Features (Socioeconomic and Building Data)	Section 3.3
Aerial Imagery Dataset Information	Section 3.5
Overview of Image Segmentation Models (Segmenter, PSPNet, KNet)	Sections 2.3, 3.6
Process for Finetuning each Model	Section 3.7
Masking Algorithm for Aggregating Land Cover Results	Section 3.8
Qualitative and Quantitative Segmentation Results	Section 4.1
Flood Risk Model Output: High Flood Hazard Risk Census Tracts	Section 4.2
Flood Risk Model Output: High Flood Vulnerability Risk Census Tracts	Section 4.3
Flood Risk Model Output: High Risk Census Tracts (Overall)	Section 4.4
Evaluation of Overall Flood Risk Model Results	Section 4.5

A unique aspect of this flood risk model is that it takes environmental, societal, and building features into account when determining a risk score. To provide context to the rest of the paper, we define flood hazard and flood vulnerability risks. Flood hazard risk is defined as a "quantitative or qualitative assessment of classification, area and spatial distribution of flood" [23]. Therefore, flood hazard values are due to environmental and geographic factors. In contrast, flood vulnerability is determined by social and economic factors that increase the risk of damages from hazards for individuals, communities or assets [27]. Features that are important for flood vulnerability are those related to socioeconomic factors and building characteristics. In this research, we build a model that takes a weighted combination of these flood hazard and vulnerability features to assign a single risk score to each census tract in the Boston area.

For the case of modeling flood risk, we consider the following flood hazard features:

- Tree Area
- Grass Area
- Soil Area
- Building Area

- Pavement Area
- Water Area
- Elevation
- Slope
- Flow Accumulation

In addition to these features, we take features that are a proxy to flood vulnerability. More specifically, we look at the following socioeconomic features:

- Population of adults over 65
- Population of Children Under 6
- People of Color
- Female Population
- Limited English proficiency
- Population below the poverty line
- Foreign Born Population who are not US Citizens
- Disabled Individuals

We also process building vulnerability information such as the following:

- Building Height
- Building Type
- Building Age

Many of these features were collected using datasets provided by the city of Boston as well as released data from the American Community Survey (ACS). However, the information for areas of tree, grass, soil, building, pavement, and water were obtained by semantically segmenting aerial satellite images. To develop this segmentation model, we trained and evaluated three segmentation model architectures: Segmenter, KNet and PSPNet. All three models have been pretrained on the ADE20K dataset: a collection of labeled images consisting of 150 segmentation classes. However, these images are all taken from a street view angle, whereas we want our model to work for aerial view images. Additionally, the images span a much larger and diverse set of objects as well as settings than what is necessary for this project use-case. Ultimately, the model should be able to generalize its predicted segmentations to urban aerial images consisting of the six relevant classes. From our experiments, we found that the strongest performing model for the use case of segmenting urban satellite aerial imagery is the Segmenter vision transformer model on a curated dataset. This dataset is a combination of images and corresponding segmentations from the Kaggle Aerial Semantic Segmentation Drone Dataset, Nearmap Map Browser, and Apple Maps.

After collecting the land cover data via the semantic segmentation model, the remaining flood hazard data from GeoTIFF rasters, and the socioeconomic and building vulnerability from census data, we developed a model that aggregates these feature values with weights to obtain a single risk value. The weights were assigned to each value based on a literature review to understand which features contribute more to flood hazard risk. More specifically, the relative weighting of the collected features related to hazard was inspired by the feature ranking in Negese et al. (2022) [21]. Therefore, the ranking of hazard features that is used in this thesis from most to least important is as follows: Elevation, Slope, Flow Accumulation, and Land Cover. These weights can also be updated with more research in the field. Additionally, as climate change worsens over time, it is possible that the contribution of certain features towards flood risk changes.

Chapter 2

Related Works

2.1 Flood Risk Modeling Approaches

Due to the effects of climate change exacerbating the frequency and intensity of floods, urban flood risk mapping has been an area of interest in many works. For example, there are flood maps made using hydrological and hydraulic models, GIS techniques, and statistical models. However, these kinds of hydraulic and hydrological models can be difficult to interpret, and the results are often beyond the technical understanding of stakeholders and policymakers [20].

As a result, machine learning is gaining popularity as a tool to predict the probability of natural hazards since modeling based on statistical methods often make unreasonable predictions as a result of simplifying the nature of some of these events. In addition to floods, machine learning methods have been used to predict the likelihood of hazards such as wildfires, droughts, earthquakes, and landslides [22].

For example, the authors of "Urban flood risk mapping using data-driven geospatial techniques for a flood-prone case area in Iran", trained several machine learning algorithms to predict the flood hazard and vulnerability levels across a town in Iran [12]. There has also been work using Generative Adversarial Networks (GANs) to produce photorealistic flood prediction images of different areas within the country. These realistic predicted images are helpful for flood visualization and can improve the adaptive capacity of cities [19].

Machine learning architectures such as convolutional neural networks have also been employed to make meaningful conclusions about satellite aerial imagery. For example, building damage is a common effect of natural disasters, and post-disaster assessment of such casualties can be helpful for preparing for future disasters. Therefore, Wen et al. (2019) proposed an improved "Mask Region Convolutional Neural Network" (Mask R-CNN). This CNN detects rotated bounding boxes of buildings and segments the boxes from very complex backgrounds in Google Earth satellite imagery [30].

Some studies have also taken advantage of Google Street View images in addition to satellite imagery to classify building types. The GPS coordinates of the buildings in aerial imagery are obtained using GIS, after which the corresponding Google Street View image is found. Then, by fine-tuning a CNN that was pre-trained on the Places dataset, the street view images of the buildings are classified by building type [17]. The use case of a building is an important flood vulnerability feature since residential buildings are more prone to flood damage than commercial buildings.

The work in the paper "Deep learning neural networks for spatially explicit prediction of flash flood probability" has used convolutional neural networks (CNNs) as well as recurrent neural networks (RNNs) to create flood probability maps. Flood probability maps are a useful tool for city planners since it can help with land cover and use planning and mitigation strategies [22].

Our research does not make flood risk predictions based on historical data. In contrast, we use current flood hazard and vulnerability data from datasets. Additionally, we use computer vision machine learning methods to find the land cover values.

2.2 Flood Risk Models that Weight Features

There are existing works that process features related to flood risk and use methods such as the Analytical Hierarchy Process to find weights for each feature to obtain a final risk value. For example, "Flood Susceptibility Mapping on a National Scale in Slovakia Using the Analytical Hierarchy Process" creates a map of regions in Slovakia

that are susceptible to flooding. In the paper, seven flood conditioning factors are chosen: distance from rivers, river network density, flow accumulation, elevation, slope, curve number, and lithology. Using the Analytical Hierarchy Process, the results from the paper determine that slope is the most important criteria affecting flood susceptibility in a region followed by river network density, distance from rivers, flow accumulation, elevation, curve number, and lithology [29].

A similar paper, "Potential flood-prone area identification and mapping using GIS-based multi-criteria decision-making and analytical hierarchy process in Dega Damot district, northwestern Ethiopia" also weights features using the same technique for a different area in the world. This paper analyzes a larger number of features and are the following: elevation, slope, flow accumulation, distance from rivers, annual rainfall, drainage density, topographic wetness index, land use and land cover, Normalized Difference Vegetation Index (NDVI), soil type, and curvature. However, the order of features they found the most important were different from the ones in Vojtek et al. (2019). This paper found elevation to be the most important factor followed by slope, flow accumulation, annual rainfall, distance to rivers, drainage density, topographic wetness index, land use & land cover, soil types, NDVI, and curvature [21]. There exists more related works in this category, yet each of them have a different collection of flood related features that they analyze, as well as inconsistent relative weightings for the features that do overlap between the works.

2.3 Semantic Segmentation

There exists several different semantic segmentation architectures and algorithms and the field has been growing and progressing at a rapid pace. One earlier image segmentation model architecture is the Pyramid Scene Parsing Network (PSPNet) presented at CVPR 2017. PSPNet aims to tackle the challenge that CNNs face when encountering diverse scenes. This occurs because these models often fail to exploit global contextual information due to the small empirical receptive fields of CNNs. To combat this, PSPNet proposes a pyramid scene parsing network which

serves as a global contextual prior. Traditional CNNs have a global average pooling layer, which is not enough to cover the necessary information in complex scenes. The PSPNet paper introduces a pyramid pooling layer, which constructs a hierarchical global prior that consists of information from different regions and along multiple scales. The different levels of features are then concatenated to create the output of the pyramid pooling global feature [32].

A more recent image segmentation model architecture is KNet which was presented at NeurIPS 2021. KNet does both instance and semantic segmentation through the use of learnable kernels. Each kernel is randomly initialized and learns a mask for an instance or stuff class. Some of these kernels are instance kernels whereas others are semantic ones. Using these two kinds of kernels allows for panoptic segmentation —a combination of instance and semantic predictions. During the forward pass through the model, the image features are convolved using the kernels to get the segmenations. The kernel update process is done dynamically so that they are conditional to their activations. This dynamic process makes the model kernels more discriminative which leads to better segmentation performance. For assigning the target to each kernel, KNet employs a bipartite matching strategy. This is an advantageous method because it creates a one-to-one mapping between kernels and image instances, therefore avoiding cases where one kernel maps to many instances in the image [31].

In addition to CNN based methods, there have also been the rise of Vision Transformer (ViT) models for image segmentation tasks. For example, "Segmenter: Transformer for Semantic Segmentation" (ICCV 2021) is one such model that uses a vision transformer instead of CNNs. One drawback of CNN based segmentation models is that the localized nature of the kernels often miss the global information of the image. Therefore, Segmenter frames the problem of semantic segmentation as a sequence to sequence problem and uses an encoder-decoder transformer structure to map patch embeddings to semantic class annotations. First, each image is split into a sequence of patches, and each patch is flattened to be a vector. Each patch vector is then projected to obtain a sequence of patch embeddings. To learn positional information, which is

lost when splitting an image into patches, a learnable positional embedding vector is concatenated to the sequence of patches. The transformer encoder then encodes this sequence using a multiheaded self attention layer and a point-wise Multi-Layer Perceptron (MLP). This encoded information is then passed to the transformer decoder which learns to map each patch encoding to a patch level class score. The patch level scores are then upsampled to obtain a class-score at the pixel level. Although vision transformers require large datasets to train, Segmenter pre-trained on a large dataset such as ADE20K only requires a moderate sized dataset for finetuning. [28].

After reviewing the literature, we find that there has been minimal to no work to our knowledge in determining the flood risk of an area by aggregating relevant features from satellite image data through semantic image segmentation. Additionally, the combination of these image features with socioeconomic and geographic datasets as well as building data to determine the flood risk at a granular level is another novel aspect to this study. This is because the majority of flood risk mapping is done at the city level, and building features are rarely used. Another existing problem in flood risk mapping and land cover studies is that acquiring detailed land cover data from GIS systems is labor intensive and expensive [26]. Therefore, it is more difficult for poor countries to hire manpower to collect this data. In contrast, using a trained image segmentation model on aerial imagery would be much cheaper and less time consuming. To reiterate, one of the main research contributions in this thesis is an image segmentation model trained on urban aerial imagery that classifies each pixel into one of six relevant categories: building, trees, grass, pavement, soil/Earth, and water.

There is also an abundance of aerial satellite imagery in the form of Google Maps and Apple Maps. Therefore, there is plenty of available data to run the image segmentation models on. Additionally, we combine more relevant flood features than existing studies to determine the flood risk. Utilizing these features that directly influence the flood hazard and vulnerability risks can give a more clear and comprehensive picture of the flood risk of an area.

Chapter 3

Methods

3.1 Overview

The objective of this chapter is to give an in-depth description of the project methods. To reiterate, the main contribution of this research is a flood risk model that utilizes several flood hazard and vulnerability features across Boston area census tracts. The features we collect can be organized into flood hazard and flood vulnerability features.

Below we show the category that each feature falls under:

1. **Flood Hazard Features**
    (a) Urban Land-Cover
        i. Building Area
        ii. Grass Area
        iii. Tree Area
        iv. Pavement Area
        v. Soil Area
        vi. Water Area
    (b) Elevation Data
    (c) Slope Data
    (d) Flow Accumulation Data

2. **Flood Vulnerability Features**

   (a) Socioeconomic Data

      i. Older adults (Over 65)

      ii. Children (Under six)

      iii. Limited English Proficiency

      iv. Individuals Below Poverty Line

      v. Disabled Individuals

      vi. Female Population

      vii. People of Color

      viii. Foreign Born Population (Non-US Citizen)

   (b) Building Data

      i. Building Height

      ii. Building Age

      iii. Building Type

To find the urban land cover of Boston, we find the areas of grass, tree, building, pavement, soil, and water from satellite aerial images of Boston. For the remaining flood hazard features, we find the average elevation, flow accumulation, and slope of each census tract using raster datasets. Along with these flood hazard related features, flood vulnerability features such as socioeconomic data and building data were also collected and processed. These particular flood hazard and vulnerability features were chosen after a literature review of factors that affect flood susceptibility. Additionally, all of these features had datasets available for the Boston region. This chapter will go into more detail over why the above features were chosen and their impact towards flood risk. Figure 3-1 displays an overview of the flood risk model and all the features that are used in the model to find the flood risk of census tracts. The aerial images of Boston are fed through a fine-tuned Segmenter ViT to obtain the percentage of buildings, trees, grass, pavement, and water for each census tract. These values are inputs to the flood risk model, which essentially weights the features.

The other relevant flood hazard data are the elevation, slope, and flow accumulation. Lastly, features related to flood vulnerability such as socioeconomic data and building data are also inputs to the flood risk model. The flood risk model weights all of the features and outputs a risk value for each census tract. These risk values are then sorted to find the census tracts that are most susceptible to flood risk.

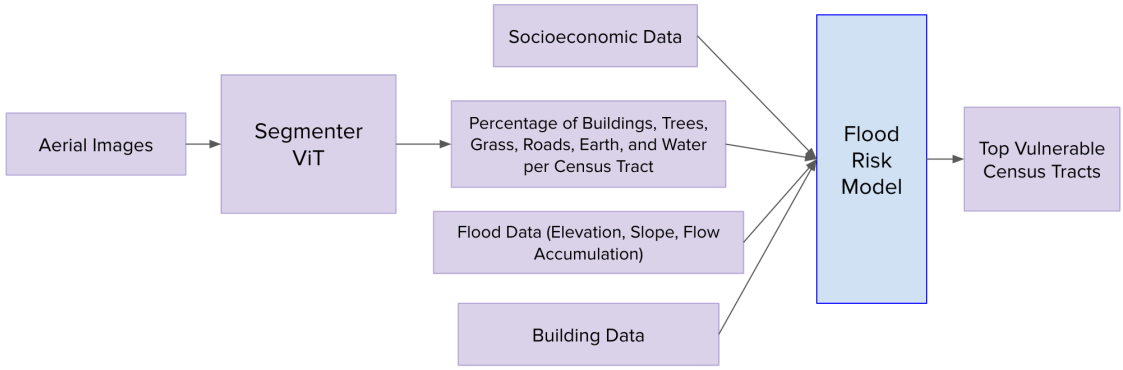

Figure 3-1: Diagram displaying the inputs and outputs of the flood risk model

3.2 Flood Hazard Features

This section outlines the features collected that relate to flood hazard. Flood Hazard is risk that is due to environmental and geographical factors and they correspond to the severity of a flood in a certain location. However, they do not provide any insight on how the people and infrastructure in an area are affected.

3.2.1 Land Cover

Urban land is characterized by large spaces of concrete roads and sidewalks as opposed to greenspaces and vegetated areas. In this section, we describe the types of land cover features we are interested in. To find the values of each feature, we trained an image segmentation model that detects the areas of each land cover type from aerial images.

Trees

According to some studies, trees planted in urban settings can mitigate runoff and flooding. In general, it is more desirable for cities to invest in green infrastructure than traditional flood mitigation devices such as barriers and drains. The study showed that cities with more trees and greenery do well in storms because trees can absorb water and also intercept them through their canopies and roots. Additionally, trees store water in their leaves and release it slowly through evaporation. As a result, urban locations with more trees see reduced runoff from storms. A study in British Columbia showed that trees are able to intercept 50 to 60 percent of rainfall [10]. Therefore, it is probable that neighborhoods and census tracts in Boston that have an abundance of trees and greenery can fare better in flood events.

Soil and Grass

Areas with soil, grassland, and other forms of vegetation are responsible for capturing a portion of rainfall as well as flood water that falls on it. By segmenting the satellite data for these areas of vegetation and soil, we find the percentage of pixels that are classified as soil and grass as opposed to the pixel percentage of concrete areas. If there is a higher percentage of pixels classified as soil and grass than concrete, the lower the flood hazard of that census tract. [15].

Concrete Spaces

As mentioned before, concrete spaces are not ideal for flood risk mitigation because of their lack of ability to absorb water. Therefore, knowing the area of concrete in each Boston census tract is valuable information for determining the flood hazard risk of that area. The fine tuned semantic segmentation model is able to identify paved and concrete areas such as roads and sidewalks.

Water

Areas that are closer to coasts and and bodies of water will be hit harder during coastal flooding events. Additionally, the levels of water bodies such as rivers increase during rainfall events. As a result, overflow into surrounding areas often occur [27]. Therefore, a census tract containing water is at a higher flood hazard risk than a census tract that is farther from bodies of water.

3.2.2 Elevation

Census tracts that are closer to sea level will face higher inundation levels due to rainfall. Additionally, lower elevation areas are also more susceptible to flooding due to the effects of gravity influencing the flow of water from higher to lower lying regions [27]. Raw elevation data was obtained from EarthENV as a GeoTIFF file at a resolution of 1 KM, and the data itself is from the Global Multi-resolution Terrain Elevation Data (GMTED2010).[1] A shapefile of the Boston census tracts was used to mask and crop the elevation raster file to find the average, minimum, maximum, and median elevation values for each census tract [2]. More specifically, the `rasterio.mask.mask` module was used to do this, and the elevation values in the raster are expressed in meters from sea-level.

3.2.3 Slope

Previous work measuring flood risk determine that the slope of an area is a key factor in flood susceptibility. Slope is the change in elevation between two areas, and locations that have low elevation and low slope values are more prone to flood risk. Low slope areas tend to store runoff accumulated from the higher slope areas. In contrast, areas with a steeper slope can dispose of runoff water due to it flowing down the land with a larger velocity [24]. Slope data was also obtained from EarthENV in the GeoTIFF format at a resolution of 1 KM. The original source of the data is the

[1] https://www.earthenv.org/topography
[2] https://data.boston.gov/dataset/census-2020-tracts

Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). Similar to the process for calculating elevation, for each Boston census tract in the shapefile, the geometry for the census tract polygon was used to mask the Slope GeoTIFF file and calculate the average, maximum, minimum, and median slope values in degrees.

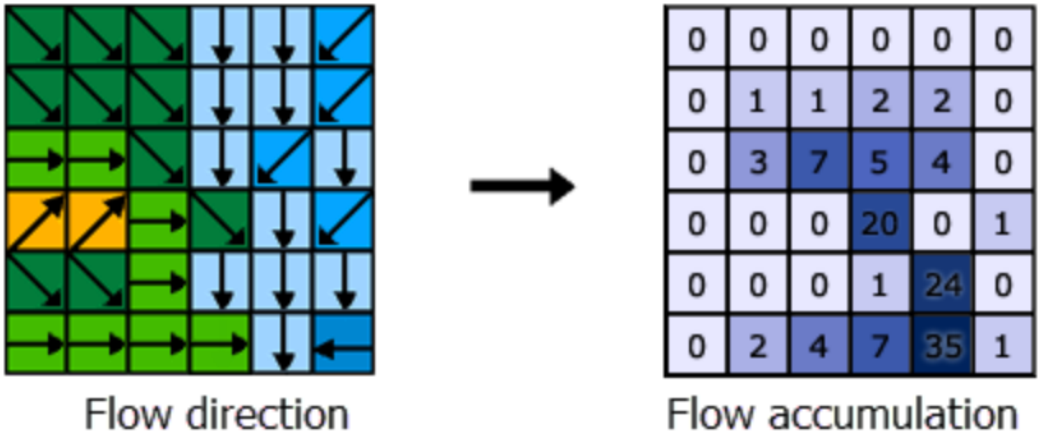

Figure 3-2: How Flow Accumulation is represented in a raster file

3.2.4 Flow Accumulation

Rainfall and precipitation are relevant features that contribute to flooding. However, it was difficult to find precipitation or rainfall datasets at the census tract level because most rainfall data is reported at the city level. In contrast, flow accumulation is a measure that is better suited as a feature for flood risk. Flow accumulation is an algorithm that determines the upstream contributing cells to each cell in a raster. It is a measure of how many cells will drain into a given cell shown in 3-2. Flow accumulation has also been used as a feature in other flood risk models and an increase in flow accumulation reflects an increase in flood susceptibility [29]. The raw data was in the format of a GeoTIFF file obtained from HydroSHEDS [3]. To find the flow accumulation for each census tract, the GeoTIFF file was masked and cropped with the shapefile containing the Boston census tracts. For each census tract, the average,

[3] https://www.hydrosheds.org/hydrosheds-core-downloads

maximum, minimum, and median flow accumulation values were calculated in number of cells.

3.2.5 Weighting Flood Hazard Features

Before weighting the flood hazard features, we normalize the values to be between 0 and 1. Then, we use weight values that mirror those used in other flood model literature. For example, the study "Flood Susceptibility Mapping on a National Scale in Slovakia Using the Analytical Hierarchy Process" uses analytical hierarchy process to find the weights of their flood specific features. They find slope degree to be the most important feature, followed by river network density, distance from rivers, flow accumulation, elevation, curve number, and lithology [29].

There are other studies as well that use analytical hierarchy process to find the weights for features relevant to flood risk. However, there is no consistent weighting of features, and the ordering of feature importance also varies. For example in "Potential food-prone area identifcation and mapping using GIS-based multi-criteria decision-making and analytical hierarchy process in Dega Damot district, northwestern Ethiopia", the order of features is the following: Elevation, Slope, Flow Accumulation, Annual Rainfall, Distance to Rivers, Drainage Density, Topographic Wetness Index, Land Use and Land Cover, Soil Types, NDVI, and Curvature [21]. For this research, we pick weights that reflect those in this paper. This is because the features we processed are a subset of the features that are weighted in that study.

3.3 Flood Vulnerability Features

In contrast to flood hazard, flood vulnerability is a measure of how the same flood event affects different groups or areas. For example, areas with a higher population of marginalized individuals or places that have weak building structures, have higher flood vulnerability levels. This section details the different flood vulnerability features processed in this research and why they are important.

Building Feature	Possible Values				
Building Type	Industrial Residential (Mixed-Use)	Residential	Vacant	Commercial	Public & Quasi-Public
Building Height	[1, 2) floors	[2, 6) floors	[6, 10) floors	10+ floors	
Building Age	Pre-1915	1915-1949	1950 - 1979	1980 - Present	

Table 3.1: Possible values that each of the building features can take

3.3.1 Building Features

As mentioned in the introduction, a climate vulnerability assessment of Boston found that over 11,000 structures will be adversely affected by stormwater flooding as a result of rising sea levels and an increase in precipitation [3]. Therefore, a large indicator of the flood vulnerability risk of a census tract is the type of buildings that exist there. This section discusses the building information we processed as well as how they specifically contribute to flood vulnerability risk. All of this building information was obtained from the Boston Buildings Inventory Dataset [2] which is in the form of a CSV file. We processed the dataset for three features: building height, building type, and building age. The possible categories that each building feature falls under is listed in Table 3.1.

Building Height

If there is a higher density of taller buildings in a location, then the individuals in those buildings can seek refuge at higher floors in the event of a flood. It is difficult for people to evacuate their homes during a flood to find shelters or buildings. This is especially true in the case of flash floods which can occur within 3 hours of heavy rainfall. Therefore, the height of buildings is a relevant feature that gives insight to the vulnerability of a census tract to flooding. The Boston Buildings Inventory dataset contains a mapping between the building ID number to the census tract number it belongs to as well as the number of floors the building has. To process this data, we decided to divide the number of floors into 4 buckets: 1 floor, 2 to 5 floors, 6 to 9 floors, and greater than 10 floors. Then, we iterated through the buildings and placed each in their corresponding buckets. We also kept a count of the total number of buildings in each census tract so that the counts in each height bucket could be

converted to percentages.

Building Type

If there are more residential buildings in an area, there is a higher flood vulnerability risk. The City of Boston's Climate vulnerability assessment says that the "majority of loss expected throughout the city will be to residential and mixed-use properties". This is especially concerning when these buildings house over 88,000 people [3]. Additionally, buildings such as universities, government buildings, fire stations, police stations, malls, and office buildings are more resistant to flood related damages. This is because these structures are often larger, and their foundations are usually deeper in the Earth. Additionally, they are more sophisticated and have more expensive mechanical, electrical, and plumbing systems that are more resistant to damage from natural disasters. The Boston Buildings Inventory dataset categorizes each building as one of the following six use-cases: Public & Quasi-Public, Residential, Residential (Mixed-Use), Vacant, Commercial, and Industrial. Similar to processing the building height information, we aggregated the percentage of buildings in each census tract belonging to each of the six categories.

Building Age

A study of flood risk in three neighborhoods in Bangladesh has shown that there is a statistically significant relationship between the age of a building and flood damage [14]. The study shows that 64.3% of buildings in the area that were between 0 and 5 years old had low damage after a flood event in contrast to 7.7% of buildings between 16 to 20 years. Therefore, a census tract that has a higher proportion of older buildings is more likely to be vulnerable to flood risk than a census tract with newer buildings. The Boston Buildings Inventory dataset has a record of the age of buildings and sorts each into one of the following 4 groups: Pre-1915, 1915-1949, 1950-1979, and 1989-Present. We grouped the buildings in the same census tract and found the percentage of buildings in that census tract belonging to each of the 4 buckets [2].

3.3.2 Social Vulnerabilities

In addition to the types of buildings in a census tract, there are also many social factors that influence the flood vulnerability risk of an area. These include, older adults, population with limited English proficiency, people below the poverty line, disabled individuals, female population, people of color, foreign born population who are not US Citizens, and children under the age of six. Here we explain why these socioeconomic features are necessary to collect to evaluate the flood vulnerability risk of an area. These results were obtained by processing data collected by the American Community Survey (ACS) in 2020.

- **Older adults (Over 65)** are more prone to physical vulnerabilities. As a result, it is difficult for them to evacuate in the case of natural disasters such as floods [6].

- **Limited English Proficiency:** If someone struggles to speak or understand English, then they might not be aware of news of natural disasters occurring. Therefore, it may be too late for them to prepare. Having limited proficiency in English may also make it difficult to communicate with emergency personnel [6].

- **Low to No Income:** If a family has limited financial resources, it can be difficult to prepare for floods or help out neighbors and/or friends [6].

- **Disabled Individuals:** It can be difficult for people with disabilities to prepare for natural disasters, including moving to a safer location. They are also more likely to face discrimination in times of crisis such as a flood event [6].

- **People of Color:** Areas with a greater population of people of color are more vulnerable to flood risk. Majority of Boston's population are people of color (53%). Statistically, people of color have lower levels of income than average. Additionally, they are more likely to live in areas that are affected by flood damage due to policies like redlining [6, 7].

- **Female Population:** The paper "Flood-Induced Vulnerabilities and Problems Encountered by Women in Northern Bangladesh" discusses how gender inequality plays a part in how women and men are affected by floods [9]. The study finds that women, especially poor and disadvantaged women, are more vulnerable to disasters in general. They often suffer more physical injuries from flood related damage and also experience harassment and violence in flood shelters. Although this study is specific to areas in Bangladesh, gender inequality exists worldwide.

- **Foreign Born Population (Not US Citizen):** People who are not US Citizens but born in a foreign country are more vulnerable because they often fall under the category of people with limited English proficiency. Additionally, immigrants who are undocumented may also be hesitant to seek help from law enforcement or emergency personnel due to fears of deportation [6].

- **Children Under Six Years Old:** Families who have very young children require more assistance and support during disasters such as flooding. Children are also more vulnerable from stress after a natural disaster [6]

We used a 2020 census dataset from the American Community Survey (ACS) to calculate these values for each census tract. We downloaded the shapefile that includes socioeconomic details at the census tract level for the state of Massachusetts.[4] To process the dataset, we filtered the shapefile to only consider census tracts in Suffolk County. The shapefile had many table headings in the format of a string code. We mapped the string code to the actual value using a reference text file and found the string codes that are relevant for each feature.[5]. For example, for population over 65, we had to aggregate the male and female populations for each age bracket over the age of 65. Similarly, for finding the total population of people who have a limited English proficiency, we had to combine the speakers of each language that did not

[4] https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-data.html

[5] https://www2.census.gov/geo/tiger/TIGER_DP/2020ACS/Metadata/TRACT_METADATA_2020.txt

speak English fluently. The number of people belonging to each vulnerable population was then divided by the area of the census tract (found using the Boston census tract shapefile) to find the population density.

3.4 Computational Resources

Training machine learning models especially larger transformer arcitechtures, require parallel computing power not capable by CPUs. Therefore, we used MIT SuperCloud in order to train the semantic segmentation models on GPU resources [25]. To train each segmentation model, we utilized 1 GPU unit and 20 CPUs.

3.5 Image Segmentation Datasets

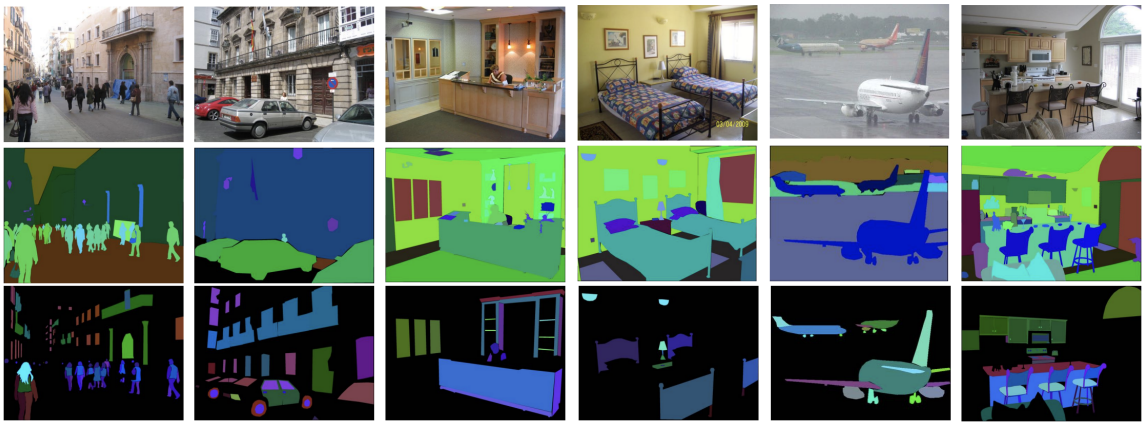

Figure 3-3: Example Images and Segmentations from the ADE20K Dataset

3.5.1 ADE20K

Every model used in this research is pretrained on the ADE20K dataset. This is a densely annotated dataset used for scene parsing. The training set consists of 20,210 images with 150 semantic classes [33]. Although the models are trained on this large dataset, finetuning was needed because the images in ADE20K are taken from a street

view perspective (as seen in Figure 3-3) whereas the view that we are interested in segmenting is aerial.

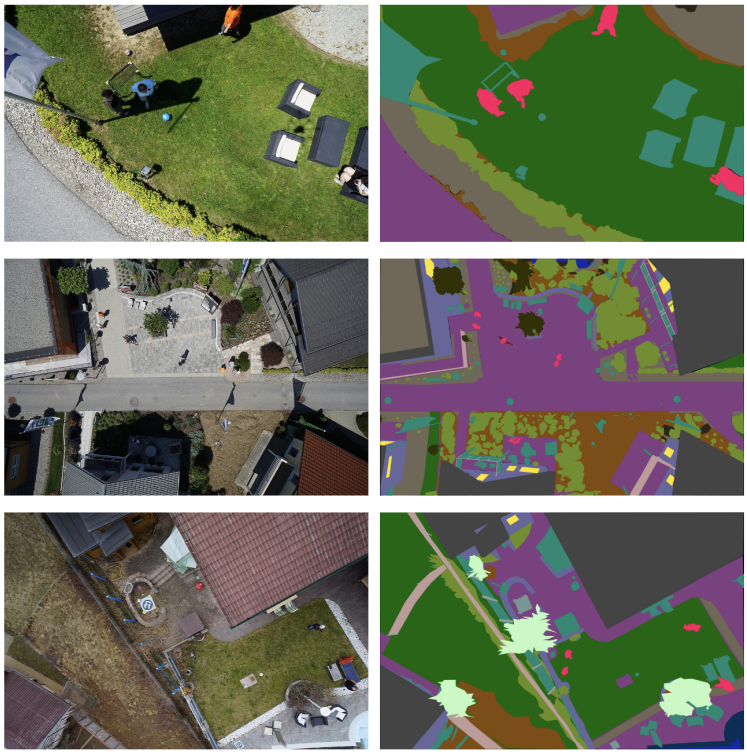

Figure 3-4: Example Images and Segmentations from the Kaggle Aerial Semantic Segmentation Drone Dataset

3.5.2 Drone Dataset

Initially, we did not have a labeled dataset of Boston aerial images to finetune the baseline models with. Therefore, to test out the finetune and evaluation pipeline, we used the "Aerial Semantic Segmentation Drone Dataset" from Kaggle [1]. We also used this dataset to observe how well the models finetune with aerial images. This dataset consists of 24 labeled classes and most of the scenes are from suburban locations as seen in Figure 3-4. Although this dataset consists of images from the correct perspective, we hypothesized that the sparse nature of these training images may not transfer well to a dataset with urban aerial scenes.

Figure 3-5: Nearmap AI Layer Features as seen on the Nearmap Map Browser. The layers on the left pane are shown in the image, where the color of the building layer is orange, vegetation layer is green, and the surfaces layer is yellow

3.5.3 Nearmap

Nearmap is an aerial imagery map company that provides frequently updated, high resolution images [6]. The data is available via a MapBrowser UI, and it also provides exportable AI layers. These AI layers are overlayed segmentations on the map and include tree overhang, asphalt, vegetation, lawn grass, water body and more. Figure 3-5 shows how the Nearmap MapBrowser can be used to view different AI layers on the map. Overall, the AI layer feature is a great resource for obtaining ground truth segmentations of aerial images in Boston. The format of the AI layers were GeoPackage files, so there was some additional processing needed to be done in order to generate a dataset of images and corresponding image segmentations for training. This process is detailed in section 3.5.6.

Figure 3-6 shows an example of images from Nearmap from March 2022 and August 2022. As evident in the image, the trees in the March image have less leaves on them, which made it difficult for the image segmentation model to detect them. As a result, we included Nearmap images from August 2022 in our training, testing, and validation datasets to finetune the segmentation models.

[6] https://www.nearmap.com/au/en

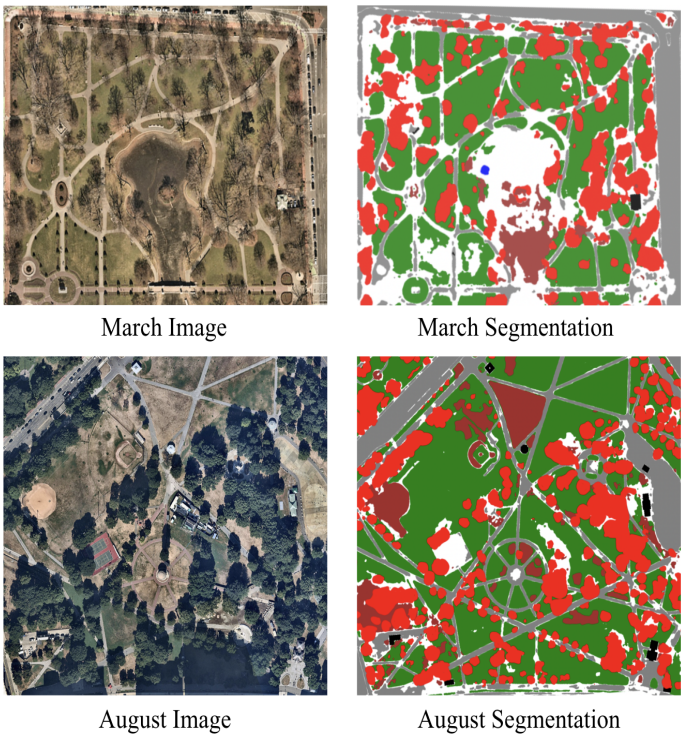

Figure 3-6: Example Images and Segmentations from Nearmap

3.5.4 Apple Maps

We noticed that Apple Maps has images with better resolution than Nearmap. Therefore, we collected some images from Apple Maps to include in our train, test, and validation dataset. However, we were able to re-use the segmentations from Nearmap with some edits. We also included additional images for which we created manual ground-truth segmentation maps for.

3.5.5 Combined Dataset

The final test and train dataset was a combination of the Drone, Nearmap August, and Apple Map datasets. The final dataset had a total of 545 images, and we used a 80%, 10%, and 10% split to have 432 train, 60 test, and 53 validation images. Each image was also resized smaller to be 600 x 400. Example images and segmentations from the combined dataset are shown in Figure 3-7.

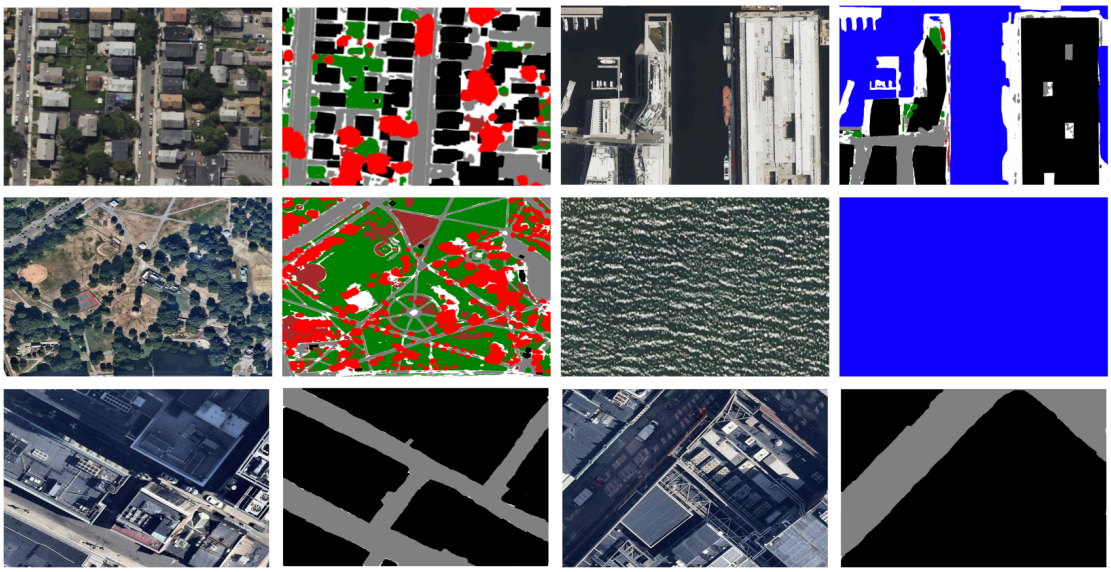

Figure 3-7: Example Images and Segmentations from Combined Dataset

3.5.6 Dataset Labeling Process

The Kaggle Aerial Semantic Segmentation Drone dataset had available annotations for each image in the dataset. However, the Nearmap dataset segmentations were given to us via GeoPackage files. Therefore, we first had to pre-process the AI layers using GeoPandas. We overlayed a visualization of each GeoPackage file for each AI layer on top of each other and assigned each a different color in the segmentation map: Building = black, Dirt/Soil = Brown, Hard Pavement = Gray, Tree = Red, Vegetation = Green, and Water = Blue. The result was a segmentation map such as the ones displayed in Figures 3-6 and 3-7. However, the segmentation map was not rotated or aligned properly to match with the corresponding aerial image. In order for an image segmentation model to train, the pixels of the image and corresponding annotation file must match up. To do this, we used Adobe Photoshop to align the segmentation with the corresponding image. After fixing the alignment, we also had to fill in some of the incomplete segmentations with Photoshop. As mentioned earlier, we had to manually create the segmentations from scratch for some of the additional Apple Maps Images. These examples are shown in the last row of Figure 3-7.

Model	Method	Backbone	Crop Size	LR Schedule	Mem (GB)	Inf Time (fps)	mIoU	mIoU(ms+flip)
Segmenter	Segmenter Mask	ViT-S_16	512x512	160000	2.03	24.80	46.19	47.85
KNet	KNet + DeepLabV3	R-50-D8	512x512	80000	7.42	12.10	45.06	46.11
PSPNet	PSPNet	R-101-D8	512x512	160000	-	-	44.39	45.35

Table 3.2: Baseline Model Details for Segmenter, KNet and PSPNet

3.6 Image Segmentation Models

To choose the image segmentation model that would extract the relevant flood hazard features from satellite aerial imagery, the performance of three different models were compared. These were PSPNet (CVPR 2017), Segmenter (ICCV 2021), and KNet (NeurIPS 2021). The task of segmenting urban aerial imagery is challenging due to the density of objects in cities. Therefore, we wanted to see how each model would fare in this specific use case. All three models are available in OpenMMLab's MMSegmentation repository [7] [11]. Additionally, all baseline models were pretrained on the ADE20K dataset and Table 3.2 shows more details of each baseline model we used.

3.6.1 Segmenter

Segmenter is an image segmentation model that has a Vision Transformer as its backbone. This was a model that we wanted to try because the filters of traditional CNN based segmentation models often miss the global information of the image due to their localized nature. Therefore, Segmenter frames the problem of semantic segmentation as a sequence to sequence problem. Additionally, pre-trained vision transformers such as Segmenter on a large dataset such as ADE20K only require a moderate sized dataset for finetuning. This is an advantageous property because annotating the satellite images of urban areas can be time consuming due to the dense nature of cities. We use one of the models pre-trained on the ADE20K dataset and picked the transformer mask decoder instead of the linear one due to better performance on the ADE20K dataset.

[7] https://github.com/open-mmlab/mmsegmentation

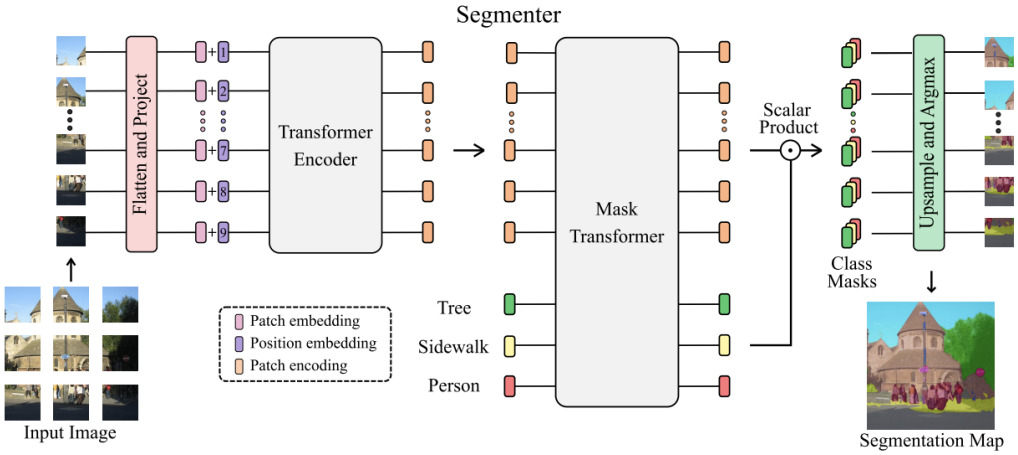

Figure 3-8: Segmenter Architecture [28]

3.6.2 KNet

KNet is another image segmentation model architecture introduced during NeurIPS 2021. Out of the three models we used, this one is the most recent. KNet is capable of doing both instance and semantic segmentation through the use of learnable kernels. In our case, we want to do semantic segmentation, and we were interested in the dynamic kernel update process that allows for more discriminative kernels [31]. The details of the KNet model we are using are displayed in Table 3.2.

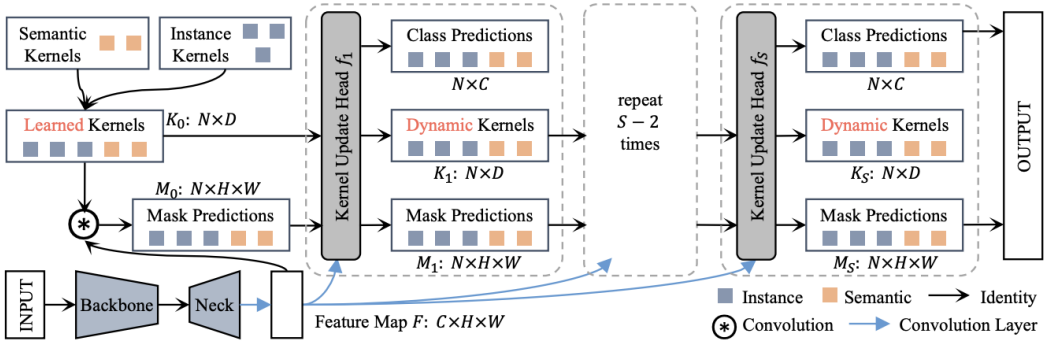

Figure 3-9: KNet Architecture [31]

3.6.3 PSPNet

PSPNet is an older segmentation model we were interested in. PSPNet introduces a pyramid pooling layer, which constructs a hierarchical global prior that consists of information from different regions and along multiple scales. This improves the ability of the model to capture global contextual information. We were interested in the novel pooling layer because urban areas often have uniform layouts, so the capability of PSPNet of taking advantage of the global context of an image when making segmentation predictions may be useful.

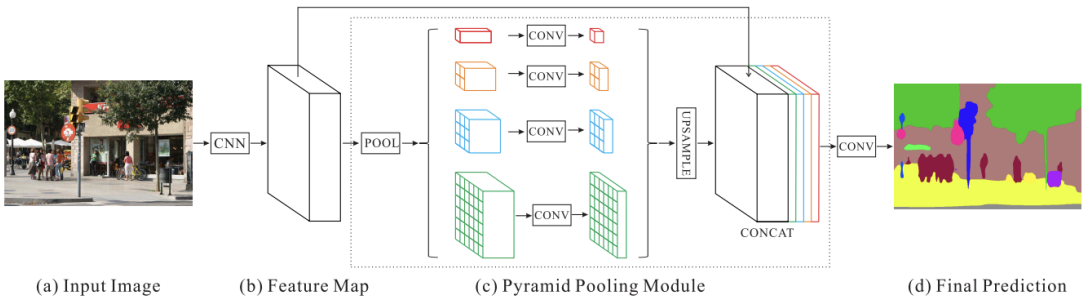

Figure 3-10: PSPNet Architecture [33]

3.7 Training Segmentation Models

Figure 3-11 shows the pipeline for training and testing each model to eventually choose one for the task of calculating land cover values in Boston. As mentioned in 3.5.5, the dataset we used to train and test the model contains images from the Kaggle Drone dataset, Nearmap, and Apple Maps as well as their corresponding segmentations. Then, after obtaining the quantitative and qualitative results, we chose the model that performs the best for the task of segmenting urban aerial imagery into different land cover classes.

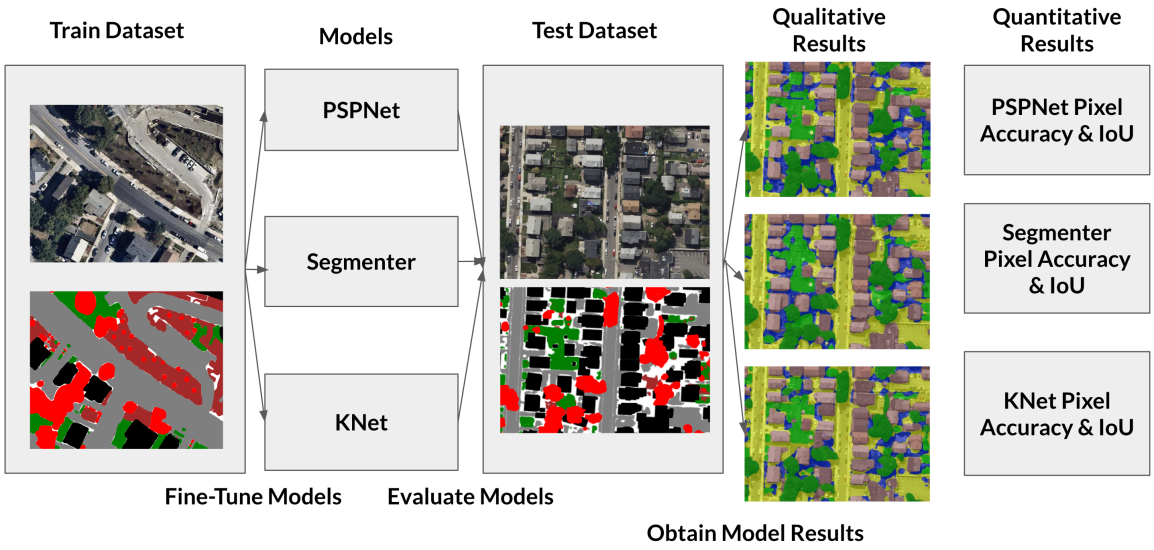

Figure 3-11: Train and Testing Pipeline to Find the Optimal Model for the Land Cover Task

3.7.1 Pre-Processing

All models used were pretrained on the ADE20K dataset. The ADE20K dataset has its own segmentation color palette as well as class IDs representing different objects. Additionally, there are 150 segmentation classes in ADE20K, whereas the Drone Dataset has 24, and Nearmap and Apple both have 6. Since ADE20K is such a vast dataset, there were corresponding class names in the ADE20K dataset for each class in the Drone, Nearmap and Apple datasets. Therefore, we wrote pre-processing code to convert the segmentation map of the Drone, Nearmap and Apple datasets to have pixel values corresponding to the class number in the ADE20K dataset. In addition to this, the images also had to be resized to a smaller size as well as converted to JPG for the segmentation models. After the images were processed, we split the image and annotation pairs into train, test, and validation sets.

3.7.2 Experiments and Metrics

We ran several experiments to find the optimal model for our task. The results from these experiments are outlined in Chapter 4 in detail. We finetuned each model on a dataset comprising of images from the Kaggle Drone Dataset, Nearmap Images,

and images from Apple Maps. As mentioned earlier, the models that we finetuned were Segmenter, KNet, and PSPNet. Before finetuning on a dataset, we initially evaluated the performance of the baseline model. Then, after finetuning, we compared the qualitatative results and quantitative results. The quantitative metric we were interested in is the Intersection over Union. We found this to be a more valuable metric over metrics such as pixel accuracy. This is because models can yield high pixel accuracy without having good qualitative results.

3.8 Calculating Land Cover

Semantically segmenting the aerial images of Boston Census tracts are used to calculate the land cover. This section details how the results from the segmentation model are aggregated to get a percentage of each type of land area (Building, Tree, Grass, Soil, Sidewalk/Concrete, and Water).

3.8.1 Aggregating Image Features Across Census Tracts

Obtaining Georeferenced Images

To find the area of trees, grass, concrete, soil, water, and buildings in each census tract, we needed to ensure that we collect a complete aerial imagery dataset that includes all areas of each census tract and unseen by the model. Simply taking screenshots from Apple Maps or Nearmaps would not be enough because we needed a way to associate each image, more specifically each pixel in the image, with a geographic location (i.e census tract). In order to associate image pixels with their world coordinate, we exported several georeferenced images from Nearmap so that together all of Boston is covered.

Georeferenced images are images that have an associated coordinate file consisting of the scale of the image per pixel as well as the (X, Y) world coordinates of the top left pixel in the image. An example of how this is done is shown in Figure 3-12. In the figure, the red lines represent census tract boundaries from the shapefile and the white-

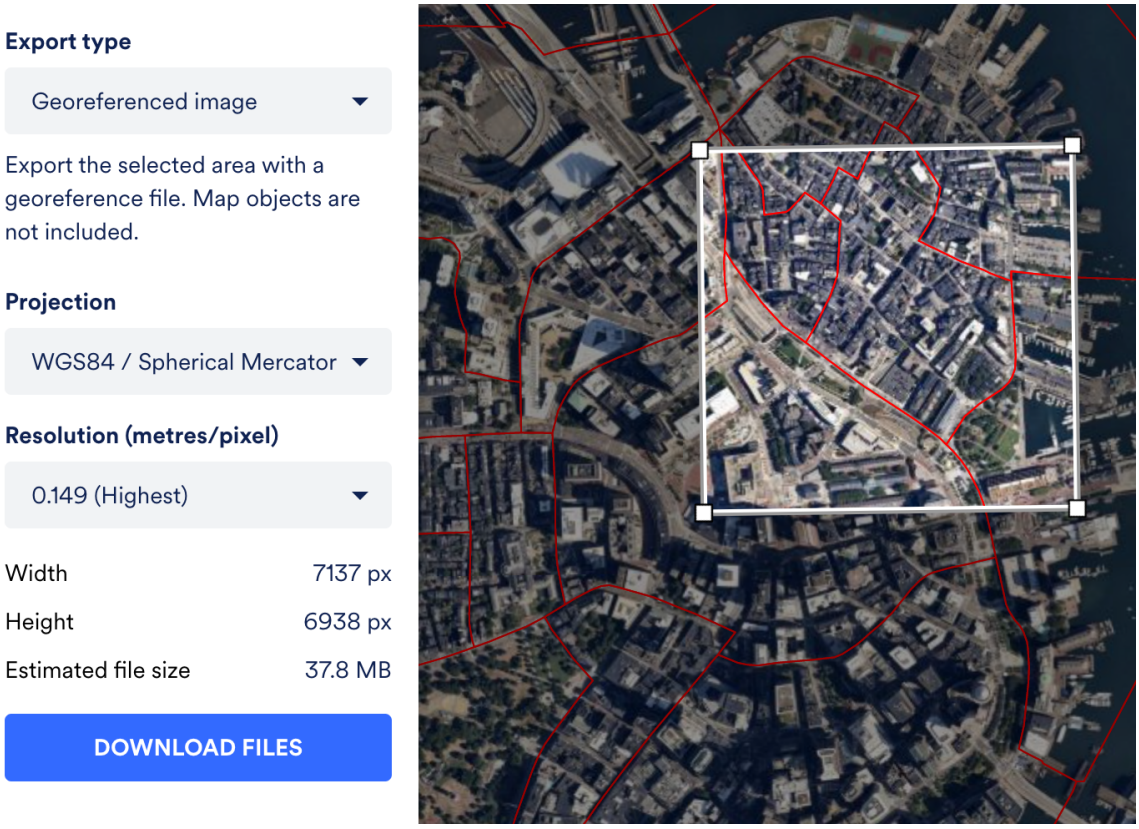

Figure 3-12: Exporting Georeferenced Images from Nearmap Map Browser. The coordinate projection is in WGS84. The red outlines on the map separate each Boston census tract

bordered rectangle is the area for which the image and its associated coordinate file is downloaded. The resolution at which we download the image is at 0.149 meters/pixel. Additionally, the dimensions of each image was approximately 7000 pixels wide and 7000 pixels high. As seen in the image, a single image can contain several census tracts within it. Therefore, we produced masks for each image that has a white pixel in the area that corresponds to belonging to a given census tract, and the rest of the image having black pixels. As a result, when we obtained the image segmentations, we knew which pixels to count in each census tract to avoid over/undercounting.

Using the georeferenced images, we converted the pixel coordinates (x_{pixel}, y_{pixel}) to their world coordinates using the following equations. In the equations, x_S is the pixel width, y_S is the pixel height, and (X_{TL}, Y_{TL}) is the world coordinate of the top left pixel in the image.

$$X_{world} = x_S \cdot x_{pixel} + X_{TL} \tag{3.1}$$

$$Y_{world} = y_S \cdot y_{pixel} + Y_{TL} \tag{3.2}$$

Creating Masks for each Image and Census Tract

In addition to the georeferenced images, we also used a shapefile that consists of the 2020 Boston census tracts. This is the shapefile that is displayed in Figure 3-12 and it consists of the world coordinates for each census tract boundary. Our approach was the following: We iterated through each census tract in the shapefile and checked if there is an intersection between the given census tract geometry and the polygon of the georeferenced image. We were able to do this by converting the corner pixel coordinates of the image to their world coordinates and creating a Shapely Polygon. If there was no intersection, we moved on to the next image. Otherwise, we found the intersecting pixels by using the `skimage.measure.grid_points_in_poly` function and created the mask for that census tract and image pair. Note: One image can have multiple masks (since an image can contain more than one census tract).

After we generated the masks, we processed the images some more. Image inputs to each segmentation model must be resized smaller, yet the georeferenced images and their associated masks were very large. However, simply resizing the images would have caused a problem with the coordinate file specifications (since the dimensions of the image would have changed). Additionally, there may have also been a problem with the location of mask pixel values when resizing the mask smaller. To avoid these problems, we cropped each image and mask, and ran the segmentation model on the cropped images.

Applying Masks to Segmentations

After obtaining the segmentations from running the model on the cropped images, we masked each image with their associated masks (the number of masks for each

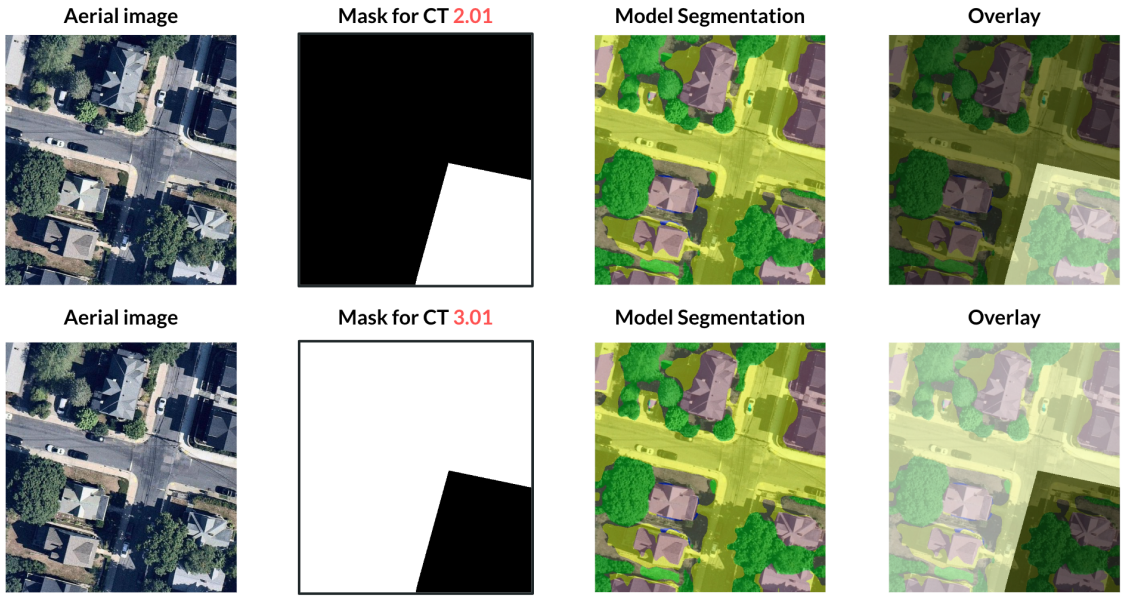

Figure 3-13: Masking Process

image is equal to the number of census tracts represented in that area). For example, the aerial image in Figure 3-13 has two census tracts in it: 2.01 and 3.01. Therefore, there are two masks that are inverses of each other. The masks are overlayed on the segmentation to only consider the pixels in the white areas of the mask to be added to the land cover count for that census tract.

Chapter 4

Results

4.1 Semantic Segmentation

The performance of each baseline model is displayed in Table 4.1, the quantitative results from the finetuned semantic segmentation models are displayed in Table 4.2, and the qualitative results are in Figure 4-2. The results displayed here are from evaluating the models finetuned on the combined dataset consisting of images from the Kaggle Drone Dataset, Nearmap, and Apple Maps.

	Baseline PSPNet		**Baseline Segmenter**		**Baseline KNet**	
	IoU	**Accuracy**	**IoU**	**Accuracy**	**IoU**	**Accuracy**
Building	21.37	42.09	36.36	81.86	25.4	55.96
Tree	24.52	64.31	42.39	75.19	35.45	72.25
Grass	22.48	26.37	36.66	41.37	26.85	31.53
Sidewalk	0.91	0.91	13.47	13.93	2.65	2.66
Earth	9.18	20.86	10.99	32.61	9.41	24.6
Water	0.11	0.12	12.37	12.42	0	0
Clutter	0	0	0	0	0	0

Table 4.1: Quantitative Performance of each Baseline model on the Test Dataset

4.1.1 Performance

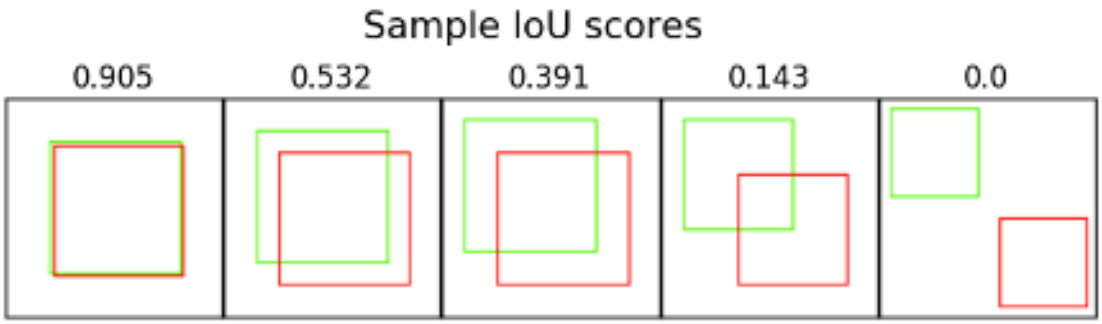

Figure 4-1: IoU Samples [8]

Metrics

The quantitative metrics that we used to evaluate the performance of PSPNet, KNet, and Segmenter on the test dataset are pixel accuracy and Intersection Over Union (IoU). Pixel accuracy is the percentage of pixels correctly classified for each segmentation class, and the IoU is a measure of how well the predicted and ground-truth segmentation maps for each class overlap each other. The value for the IoU metric ranges from 0 to 1, where 1 is a perfect overlap and 0 is no overlap between the ground truth and predicted segmentations. Figure 4-1 shows example IoU scores for each level of bounding box overlaps. An IoU score of greater than or equal to 0.532 for each class represents a segmentation model that is satisfactory for the purposes of detecting urban land cover.

Quantitative and Qualitative Performance on Test Dataset

The performance of each baseline model is displayed in Table 4.1. The poor results are expected because the baseline models were trained on ADE20K, which is a dataset with images taken from the street level. Additionally, many of the images in the dataset are not of urban or outdoor scenes. As a result, the baseline models did not generalize to the aerial images in the test set. Out of all the models, the baseline Segmenter model performs the best on the test dataset.

	Finetuned PSPNet		Finetuned Segmenter		Finetuned KNet	
	IoU	**Accuracy**	**IoU**	**Accuracy**	**IoU**	**Accuracy**
Building	65.55	81.12	71.88	85.17	67.12	82.16
Tree	61.16	78.93	65.73	82.49	65.28	79.81
Grass	86.35	93.13	86.69	92.81	87.27	92.09
Sidewalk	72.72	89.12	78.24	89.94	78.09	92.45
Earth	66.66	76.78	70.2	76.35	71.93	79.27
Water	81.49	81.98	95.78	97.85	95.12	95.84
Clutter	21.72	38.01	20.71	34.74	20.52	29.41

Table 4.2: Quantitative Performance of each Finetuned Segmentation Model on the test dataset

The performance of all three models on the test set improved significantly after finetuning. Based on Table 4.2, Segmenter outperforms the other two models for all categories except for grass and earth, for which KNet performs better in. We also include a separate category called "Clutter" which is what the model predicts if it is not able to fit the pixel into one of the six categories. The model also tends to classify heavily shadowed or occluded areas to be clutter.

Based on the qualitative results in Figure 4-2, Segmenter outperforms KNet and PSPNet for the example images. All three models do well at classifying trees and grass. However, for the first example aerial image, Segmenter classifies less of the grass area to be clutter (the areas in blue) and also does a better job classifying buildings. Although KNet and PSPNet do not perform as well for segmenting general and larger areas of the image, they can segment more finegrained areas of the image. This is observed in the first example where KNet is able to distinguish the small buildings surrounded by grass better than Segmenter.

In the second example in Figure 4-2, Segmenter produces better segmentation predictions than KNet and PSPNet because it correctly classifies the areas of water. On the other hand, KNet finds the large portion of the water as clutter, whereas PSPNet predicts it to be pavement. Lastly, all three models correctly classify the

trees and grass regions in the third image example comparably. However, Segmenter predicts less areas as clutter and also correctly classifies more paved areas than the other two models. Additionally, KNet classifies the boundaries of the trees to be soil/Earth even though the ground truth segmentations do not have those areas labeled as soil/Earth. Due to Segmenter having better quantitative and qualitative results, we used this finetuned Segmenter model as the land cover feature extractor on Boston aerial images.

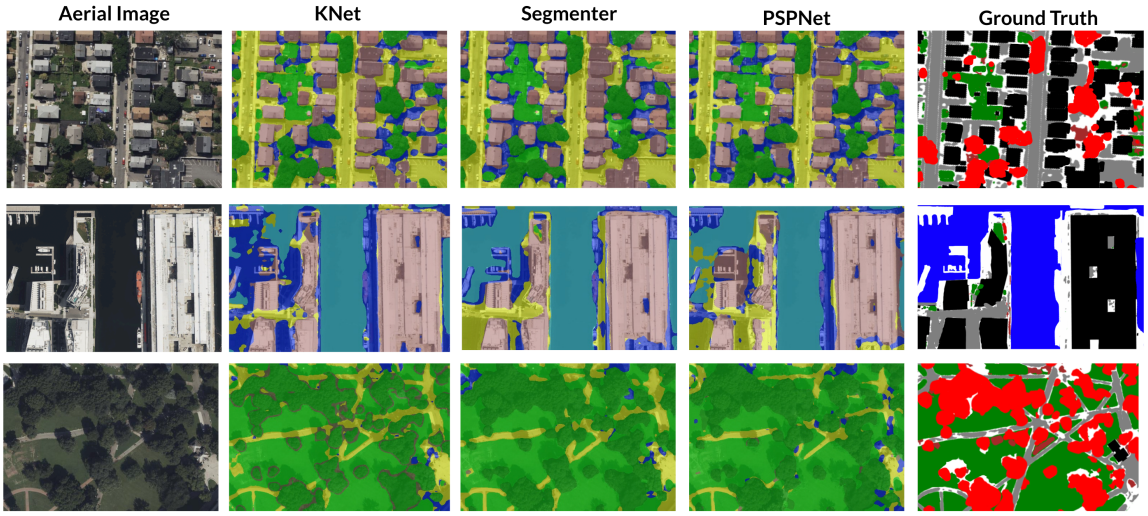

Figure 4-2: Example Qualitative Results after Evaluating Fine-tuned Models on Test Dataset

Qualitative Performance on Boston Aerial Images

Figure 4-3 shows a few examples of the performance of each finetuned model on a few examples of the Boston Aerial Images that we collected from Nearmap. These are the images that we used to find the areas of building, grass, tree, paved area, soil/earth, and water for each census tract in Boston. Therefore, these segmentations cannot be quantitatively evaluated. For the first example image, Segmenter correctly classifies all of the buildings, whereas KNet and PSPNet have less complete segmentations for them. This behavior is also evident in the second and third images. These qualitative examples further support that the finetuned Segmenter model is the best choice for the given use-case in this research.

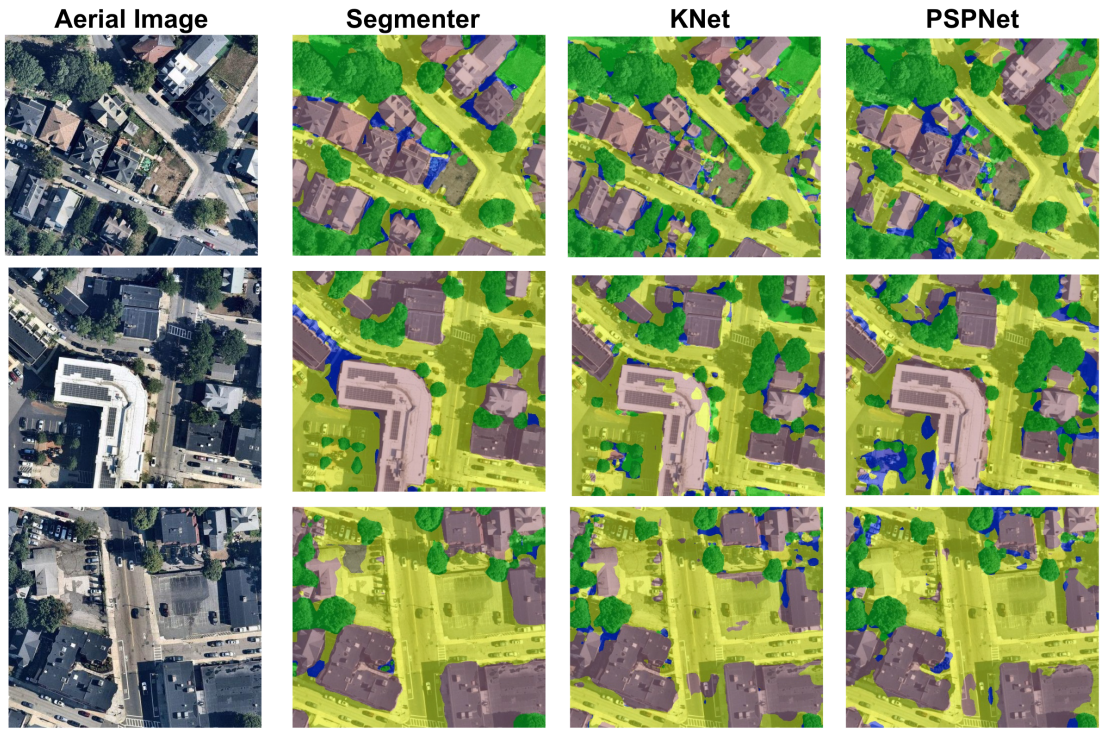

Figure 4-3: Example Segmentations of Boston Aerial Images Produced by Each Fine-tuned Model

Qualitative Performance on NYC

Although the model was trained and tested on a dataset consisting of aerial images from Boston, the model was also tested on aerial images from New York City to examine how well the model generalizes to other urban areas. The qualitative results for the Manhattan, Queens, and Brooklyn boroughs are shown in Figure 4-4. Figure 4-4 image (a) and 4-4 (c) are examples from Manhattan. Manhattan is much more urban than Boston, Queens, and Brooklyn. Additionally, Manhattan consists of a higher density of buildings that are also taller than buildings in Boston. As a result, many areas in the Manhattan aerial images are occluded by shadows. The rooftops of Manhattan buildings are also more complex than those of Boston, Queens, and Brooklyn. However, the qualitative results from the model are reasonable, and sub-figures (a) and (c) in Figure 4-4 give a few examples of how the model is able to infer the heavily shadowed areas correctly.

Pairs (b) and (d) are images from Queens and (e) and (f) are from Brooklyn.

In contrast to the Manhattan images, these are less urban and are similar to some areas in Boston. As a result, the model performs reasonably well on these images. As noticed in the performance of the Boston aerial images as well, the model still has difficulty classifying heavily shadowed areas although it had improved after making the training dataset more comprehensive.

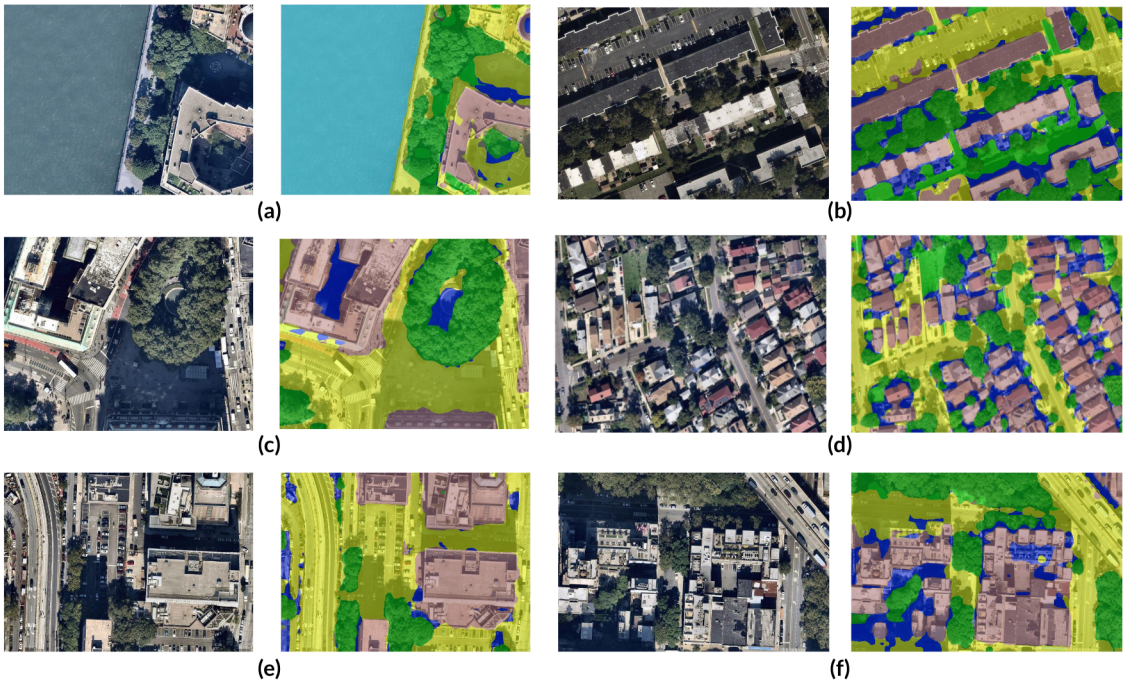

Figure 4-4: Example Qualitative Results for NYC Boroughs

4.2 Flood Hazard Features

4.2.1 Land Cover Results

Top 5 Grass		Top 5 Tree		Top 5 Earth		Top 5 Paved		Top 5 Water		Top 5 Building	
CT	Percentage	CT	Percentage	CT	Percentage	CT	Percentage	CT	Percentage	CT	Percentage
9816.0	22.41	9807	54.32	9815.02	21.52	9813	43.80	9812.01	43.91	701.02	31.43
9809	22.22	9803	45.45	9816	16.37	907	42.08	9801.01	41.05	106	30.71
9811	22.09	1106.01	45.30	9813	13.50	612.02	41.79	606.04	40.95	701.04	29.71
9817	18.54	9810	43.37	9812.01	13.14	606.02	40.63	305	39.07	403	28.68
9810	17.32	1201.05	40.99	909.01	12.30	606.03	39.40	503	36.03	303.02	27.79

Table 4.3: Top 5 Boston Census Tracts with Highest Percentage of Each Class

Figure 4-5: Boston Census Tracts with Highest Percentage of Each Class

We applied the finetuned Segmenter model on 20972 aerial images of Boston that cover all census tracts. Then, we aggregated the pixel values for each class using the masking method outlined in Chapter 3, and found the top 5 census tracts with the highest percentage of each class (shown in Table 4.3). To verify that the results make sense, example aerial images for the top census tract for each class is displayed in Figure 4-5.

The census tract with the highest percentage of grass is 9816 which consists of the Belle Isle Marsh Reservation. The Belle Isle Marsh Reservation is Boston's only salt marsh area and is a nature reservation. From the aerial image in Figure 4-5, it is evident that much of the area consists of grass, meaning that the segmentation model correctly classifies the large area of grassland in the census tract.

Similarly, census tract 9807 consists of the highest percentage of trees. This census tract contains Stony Brook Park, which is a forest in Boston. Figure 4-5 displays an

image of the aerial view of the park, showing an abundance of trees. The segmentation model found census tract 9815.02 to be the area with the most Earth/soil locations. This census tract is located in East Boston and the image in Figure 4-5 is of a housing development that is located next to the Suffolk Downs Horse Racecourse. As evident in the image, there are a lot of areas with soil and dirt.

The area with the highest percentage of concrete/sidewalk is 9813 which only contains the Boston Logan Airport. It makes sense that there is a lot of concrete areas due to the presence of airplane runways. The census tract containing the largest percentage of water is 9812.01, which is the coast containing Castle Island. The entire census tract is virtually all water, therefore, this result makes sense. Lastly, the census tract with the most buildings is in Downtown Boston.

We also calculated a weighted combination of all the land cover results together to find a risk score. This risk score does not consider other flood hazard factors such as elevation, slope, flow accumulation. It is only a measure of flood risk due to land cover. The weights and results are shown in Table 4.4. All of the high flood risk locations are those that are by a body of water and also lack vegetation. These are areas such as the Seaport Harbor and the Seaport Cruise Terminal. On the other hand, the lowest risk areas in Boston are those that have a lot of vegetation, minimal concrete, and no large water bodies surrounding them.

High Flood Risk (Land Cover Only)			**Low Flood Risk (Land Cover Only)**		
Census Tract	**Location**	**Weighted Risk Value**	**Census Tract**	**Location**	**Weighted Risk Value**
606.04	Seaport	0.843	9807	Stony Brook Reservation	-0.649
9812.02	Seaport (Cruise Terminal)	0.825	9803	Franklin Park Zoo	-0.590
305	Northend	0.770	9810	Arnold Arboretum	-0.556
408.01	Charlestown	0.741	9809	Millennium Park	-0.536
303.01	Waterfront	0.731	9811	Mount Hope Cemetery	-0.483

Class	**Weight Value**
Building	1
Tree	-1
Grass	-1
Concrete	1
Earth	-1
Water	1.5

Table 4.4: High and Low Flood Risk Census Tracts Base on Land Cover Results Only

Census Tract	Location	Elevation (m)
9801.01	Boston Harbor Islands	0.411
512	Jeffries Point	3.000
503	East Boston (Facing Boston Harbor)	3.250
504	East Boston	3.500
9813	Boston Logan Airport	3.542

Table 4.5: High Risk Census Tracts due to Low Elevation

4.2.2 Elevation Results

In Table 4.5, we have the top 5 census tracts with the lowest elevation. These are the census tracts that have the highest flood risk as a result of being closer to sea level. All of these locations directly border a water body. The census tract with the lowest average elevation is 9801.01, which includes the Boston Harbor Islands such as Fort Warren, Spectacle Island, Deer Island, and Thompson Island. Census tract 9801.01 also contains areas of water surrounding the islands, as a result, the value of the minimum elevation in that area is 0 m (sea-level) making the average elevation very low (around 0.411 meters).

4.2.3 Slope Results

Table 4.6 consists of the 5 census tracts that are most susceptible to flooding based on slope data only. Lower sloped areas are prone to flooding due to runoff water remaining in the area instead of flowing out of the census tract. All of the top 5 census tracts have very similar slope values. It is difficult to verify that the processed slope results are indeed correct for the low slope areas. However, the ordering makes sense because the census tracts with the highest slope are all areas in Boston that have considerable changes in terrain based on the Google Maps Terrain feature. Conversely, areas with lower slope have little to no change in terrain.

Census Tract	Location	Average Slope (Degrees)
607	Orton Field (West Broadway)	0.219
709.01	Chester Square (South End)	0.251
709.02	Chester Square (South End)	0.287
612.01	Columbus Park/Andrew Square	0.291
505	East Boston Memorial Park	0.297

Table 4.6: Top census tracts with highest flood risk due to low slope values

4.2.4 Flow Accumulation Results

Table 4.7 displays the top 6 census tracts that have the highest average accumulation from other areas. To reiterate, flow accumulation is calculated using a flow direction raster. The flow direction raster shows the direction water will flow from neighboring cells. It is a good measure of how much water will flow into a particular area in the case of a flood event. We ignore the first result since it is a census tract comprised of only the Charles River. An interesting observation is that the rest of the census tracts are all those that border bodies of water. For example, census tract 203.05 is located near the O'Brien Highway and Zakim Bridge. This is also near census tract 404.01, and both of these locations border the Charles River. Similarly, census tract 401 is in Charlestown next to the USS Constitution museum on the Boston Harbor. Census tract 8.06 is in Allston and 1.02 is in North Brighton, and both census tracts border the Charles River.

Census Tract	Location	Average Accumulation Per CT Cell (In Number of Cells)
9815.01	Charles River	18143
401	USS Constitution, Charlestown	10987
404.01	Paul Revere Park	7528
203.05	O'Brien Highway/TD Garden	7278
1.02	North Brighton	6837
8.06	Harvard Business School	6466

Table 4.7: Top census tracts with highest flood risk due to flow accumulation

4.2.5 Flood Hazard Weighting

High Flood Hazard Risk			Low Flood Hazard Risk		
Census Tract	**Location**	**Weighted Risk Value**	**Census Tract**	**Location**	**Weighted Risk Value**
9801.01	Harbor Islands	0.180	1401.06	Bordering Stony Brook Reservation	-3.426
512	Jeffries Point	0.128	1303	Bordering Stony Brook Reservation	-2.856
606.04	Seaport	0.120	1201.05	Bordering Jamaica Pond	-2.788
9812.02	South Boston Waterfront	0.047	1106.01	Allandale Woods	-2.778
305	Northend (Facing Boston Harbor)	0.034	1105.01	Bordering Bellevue Hill	-2.759
612.01	South Boston	0.014	9807	Stony Brook Reservation	-2.710
503	East Boston (Facing Boston Harbor)	-0.038	5.02	Pine Tree Preserve	-2.633
606.03	Seaport (Bordering 606.04)	-0.040	1105.02	Bordering 1105.01	-2.573
504	East Boston	-0.042	1304.04	Roxbury Latin School	-2.498
9813	Boston Logan Airport	-0.050	5.05	Brighton	-2.491

Hazard	**Weight Value**
Land Cover	0.5
Slope	-1.5
Elevation	-2
Flow Accumulation	1

Table 4.8: Census Tracts with the Highest and Lowest Susceptibility to Flood Hazard Risk

As mentioned in section 3, we based on our weighting from the following paper: "Potential food-prone area identifcation and mapping using GIS-based multi-criteria decision-making and analytical hierarchy process in Dega Damot district, northwestern Ethiopia". As a result, the order of features in terms of contributing to flood hazard were the following: Elevation, Slope, Flow Accumulation, and Land Cover.

Table 4.8 displays the top 10 census tracts in Boston that are the most and least susceptible to flood hazard risk. The table also includes the given weight for each feature. The location that is most likely to experience flooding due to its location and geographic factors is census tract 9801.01. This includes the Boston Harbor islands such as Fort Warren, Spectacle Island, Deer Island, and Thompson Island. The average elevation value of the census tract is very low. This is partly due to the fact that the census tract includes areas of water as well, which have elevation levels of zero meters. Although there are limited areas of building and concrete on the islands, they are surrounded by areas of water which increases the flood hazard risk of an area.

Census tract 512 near Jeffries Point is the flood risk with the second highest

weighted hazard risk value. The location faces the Boston Harbor and has low slope and elevation values as well. A large portion of the census tract also includes the Boston Harbor. As a result, the land cover risk value is also higher than that of the 9801.01 census tract. An aerial image of Jeffries Point reveals minimal greenspace and vegetation.

Majority of the high risk areas border a water body. For example, the two Seaport census tracts (606.04 and 606.03) face the Boston Harbor. Additionally, they have minimal greenspace or vegetation and majority of Seaport consists of high rise buildings and concrete areas. Additionally, Seaport is very close to sea-level and has lower slope. From previous publications, low elevation and low slope areas make areas prone to flooding. However, the flow accumulation value for Seaport was quite low.

In contrast, the census tract with the lowest susceptibility to flood hazard risk is 1401.06. This is a census tract that borders the Stony Brook Reservation. Although it does not have an abundance of greenspace, trees, or Earth, the slope and elevation values are very high. This makes it unlikely that the area receives runoff from higher elevated areas. Additionally, the runoff that it receives will run down to other areas due to the high slope. The next census tract is 1303 which also borders the Stony Brook Reservation. It has similar geographic characteristics to census tract 1401.06. Many of the other low risk census tracts are in the West Roxbury area and border each other. This makes logical sense because we expect there to be similar geographical features in adjacent census tracts.

Many of the low risk census tracts are high elevation and slope areas that consist of greenspace and trees. For example, the Stony Brook Reservation is a woodland state park that consists of many trees and vegetated areas. Additionally, there are minimal concrete and grass spaces. In the low risk areas where the land cover value is higher, it is offset by high elevation and slope values. For example, census tract 5.05 has relatively no greenspace and some trees, but its high elevation and slope reduce the weighted risk value.

4.3 Flood Vulnerability

4.3.1 Socioeconomic Data

High Flood Vulnerability Areas			Low Flood Vulnerability Areas		
Census Tract	**Location**	**Vulnerability Score**	**Census Tract**	**Location**	**Vulnerability Score**
705.02	South End	5.949	9812.02	South Boston Waterfront	0.001
104.03	Fenway/Near Northeastern University	5.900	9811	Forest Hills	0.002
702.02	Chinatown	5.375	9803	Franklin Park	0.028
607	South Boston	5.179	9801.01	Boston Harbor Islands	0.030
709.01	South End (Near BU Medical Campus)	5.134	606.02	Seaport	0.042
507	East Boston	4.969	606.03	Seaport	0.155
502	East Boston	4.775	8.06	Allston (Harvard Business School)	0.207
701.02	Downtown Crossing	4.551	612.01	South Boston	0.227
501.01	Eagle Hill	4.155	1006.03	Neponset	0.266
104.04	Northeastern University	4.052	1201.05	Jamaica Plain	0.277

Feature	Weight
POC population per area	1
Disabled population per area	1
Limited English proficiency population per area	1
Female population per area	1
Foreign born and not US citizen population per area	1
Children under 6 years population per area	1
Poverty status population per area	1
Male population over 65 per area	1
Female population over 65 per area	1

Table 4.9: Census Tracts Most and Least Vulnerable to Flood Risk (Based on Socioeconomic Features Only)

Table 4.9 shows the top 10 most vulnerable census tracts as a result of marginalized populations. The table also shows the different features processed and their corresponding weights in the vulnerability score. The census tract that is the most socially vulnerable is 705.02. This is an area in the South End of Boston. Based on Census Reporter, the median age in the location is 50.2, which is about 1.5 times higher than the rest of Boston and Suffolk County. Additionally, around 42.5% of the population is below the poverty line. This is more than double the rate in Boston (17.6%) and Suffolk County (17.3%) [4].

The rest of the census tracts in the top 10 besides 507 and 501.01 also have poverty rates that are above the city and county values based on Census Reporter data. For census tract 507, the poverty rate is 13.7% and for 501.01 it is 15%. However, census tract 507 has a very high population of POC, people with limited English proficiency, and children under 6 years old. Similarly, census tract 501.01 has a high population of children under 6 years old as well as POC.

The top 11 Census tracts with the lowest flood vulnerability are areas where there's limited to no population. As a result, the vulnerability value is 0. The name of the census tracts are: 9807, 9813, 9817, 9812.01, 9815.02, 9819, 9809, 9816, 9810, 612.02, and 9815.01. These are areas in Boston such as the Boston Logan Airport, Stony Brook Reservation, Boston Commons and Boston Public Garden, the Charles River, etc. Therefore, Table 4.9 shows the top 10 census tracts that have a population greater than 0. This includes areas with a younger population and low poverty rates such as Seaport.

4.3.2 Building Vulnerability

As mentioned in the methods section in Chapter 3, we also collected and processed relevant building information for flood vulnerability. The features processed were the age, height, and type of the buildings in each census tract. The most vulnerable areas included areas that had the highest percentage of buildings under 6 floors, buildings built before 1950, as well as residential buildings. All of these features were given equal weighting. Based on the results in Table 4.10, the two census tracts that are the most vulnerable for flood risk due to building features only are 602 and 603.01. These two census tracts border each other and are located in Dorchester Heights. To verify the results, we looked at Google Street View images of the area. The area is densely populated with residential buildings that are lower in height. This matches up with the processed data, which shows that 98% of the buildings are residential and less than 6 floors in height. Additionally, based on the processed data, approximately 92% of the buildings are built before 1950.

The area that has the lowest flood vulnerability score due to building characteristics is 704.02. This area is in the South End and a street view of the location shows taller, non-residential buildings. The census tract with the second lowest flood vulnerability score is 203.01 located in the West End. More specifically, this area includes the Massachusetts General Hospital, The Liberty Hotel, and the Shriner's Children's Hospital. Majority of the other buildings in the area are similar in that they are all non-residential buildings. In fact, only 10% of the buildings in the area

are residential.

High Building Vulnerability Areas			Low Building Vulnerability Areas		
Census Tract	**Location**	**Vulnerability Score**	**Census Tract**	**Location**	**Vulnerability Score**
602	Dorchester Heights	2.882	704.02	South End	0.751
603.01	Dorchester Heights	2.855	203.01	West End	0.902
1004	Ashmont	2.847	909.01	Dorchester	0.941
1105.01	West Roxbury	2.812	9811	Forest Hills	0.941
2.01	Brighton	2.806	9813	Boston Logan Airport	1.007
911	Uphams Corner	2.799	203.04	West End	1.062
1006.01	Dorchester Center	2.781	9807	Stony Brook Reservation	1.111
2.02	Brighton	2.777	9812.02	Seaport (Cruise Terminal)	1.117
501.01	Eagle Hill	2.772	808.01	Mission Hill	1.136
502	Eagle Hill	2.766	9815.02	Beachmont	1.167

Feature	Weight
Building Age	1
Building Height	1
Building Use Case (Residential Buildings)	1

Table 4.10: Census Tracts Most and Least Vulnerable to Flood Risk (Based on Building Features Only)

4.3.3 Weighting Flood Vulnerability Features

Table 4.11 displays the results after combining the normalized building and socioeconomic vulnerability weights to get a single vulnerability score. The building and socioeconomic features each have a weight of one and were added together to get the vulnerability score. The area with the highest flood vulnerability risk score is 705.02 in the South End. This is the same census tract that is the most socially vulnerable in Boston as well. The normalized socioeconomic vulnerability value for census tract 705.02 is 1.0 and the normalized building vulnerability value is 0.849. This value agrees with other sources such as Census Reporter and Google Maps. Besides the large population of people below the poverty line and limited English speakers, Google Street View images of the area also reveal many residential buildings that are under 6 stories tall.

The next census tract with the highest flood vulnerability risk score is 502. This was another census tract that was in the top 10 for the most socially vulnerable census tracts. This area is in East Boston, near the Mystic River. The census tract

High Vulnerability Areas			Low Vulnerability Areas		
Census Tract	**Location**	**Vulnerability Score**	**Census Tract**	**Location**	**Vulnerability Score**
705.02	South End	1.849	9811	Forest Hills	0.327
502	East Boston	1.762	9813	Boston Logan Airport	0.349
104.03	Fenway/Near Northeastern University	1.760	9812.02	Seaport (Cruise Terminal)	0.387
507	East Boston	1.752	9810	Arnold Arboretum	0.420
709.01	South End (Near BU Medical Campus)	1.713	9801.01	Harbor Islands	0.421
501.01	Eagle Hill	1.660	909.01	JFK Presidential Library	0.428
610	South Boston	1.530	9803	Franklin Park Zoo	0.490
301	North End	1.516	203.04	West End	0.502
104.04	Northeastern University	1.512	8.06	Allston (Near Harvard Business School)	0.517
821	Roxbury	1.500	701.04	Financial District	0.287

Feature	Weight
Building Features	1
Socioeconomic Features	1

Table 4.11: Census Tracts Most and Least Vulnerable to Flood Risk (Based on Combining Building and Socioeconomic Results)

has a normalized socioeconomic vulnerability score of 0.80 and a normalized building vulnerability score of 0.952. Similarly, Google Street View of the area reveals many residential neighborhoods comprised of older and shorter homes.

The census tract with the lowest flood vulnerability risk score is 9811. This area is in Forest Hills and has a very small population of people living there. As a result, the normalized socioeconomic vulnerability score for the area is 0.00042 and the building vulnerability score is 0.327. Based on Census Reporter, there are only 4 households living in the area, and Google Maps shows mostly non-residential buildings in the area [4]. The rest of the low flood vulnerability risk census tracts are similar in that they have a smaller population of people living there and also a small number of residential buildings. As a result, it makes sense that these areas rank low for flood vulnerability risk.

4.4 Obtaining Flood Risk

Finally, we combined the flood vulnerability (which includes socioeconomic and building vulnerability features) with the flood hazard features to obtain a singular flood risk value for each census tract. First, the hazard and vulnerability scores were normalized, and then the flood hazard values were weighted with a value of 2. This is

High Flood Risk Areas			Low Flood Risk Areas		
Census Tract	**Location**	**Total Risk Score**	**Census Tract**	**Location**	**Total Risk Score**
709.01	South End	2.588	9807	Stony Brook Reservation	0.578
705.02	South End	2.565	1401.06	Bordering Stony Brook Reservation	0.586
104.03	Fenway	2.544	1201.05	Jamaica Plain	0.795
504	East Boston	2.508	1304.04	West Roxbury	0.808
104.04	Northeastern University	2.457	9810	Arnold Arboretum	0.837
507	East Boston	2.457	1303	West Roxbury	0.839
506	East Boston	2.428	9809	Oak Hill Park	0.866
502	East Boston	2.420	1106.01	Allandale Woods	0.878
512	Jeffries Point	2.417	9803	Franklin Park Zoo	0.879
607	South Boston	2.400	5.02	Pine Tree Preserve	0.921

Feature	Weight
Vulnerability Features	1
Hazard Features	2

Table 4.12: Census Tracts Most and Least Susceptible to Flood Risk (Based on Combining Vulnerability and Hazard Values)

because although vulnerability is important, we want to find the areas that have a risk of flooding as well as vulnerable populations living in it. Increasing the weight from 2 causes the hazard values to completely overpower the vulnerability values. As a result, a weight of 2 balances the two features while still favoring higher flood hazard weights.

The results are shown in Table 4.12 and the top census tract that is most susceptible to flood risk is 709.01 which is in the South End. The normalized flood hazard value for the census tract was 0.83 and the vulnerability weight was 0.926. Based on Census Reporter, the area has 22.6% of people who are below the poverty line, which is around 1.3% times the poverty rate in Boston [4]. Additionally, based on the processed flood hazard data, census tract 709.01 has low slope and elevation values (normalized values are 0.014 and 0.137 respectively). The maximum value that each of these features can take are 1. Additionally, the normalized flood risk from land cover is relatively high (0.55) where the maximum value is 1.

From the processed flood vulnerability data, census tract 709.01 has a normalized socioeconomic vulnerability score of 0.86 and a normalized building vulnerability score of 0.85. More specifically, the normalized POC population density is 1 (maximum), the disabled population density is also 1, and the limited English proficiency

population density is around 0.93. As a result, this area is very socioeconomically vulnerable. The census tract is also vulnerable in terms of buildings. Based on the processed Boston Buildings Dataset data, 88% of buildings in the area are residential buildings and 87% of buildings are less than 6 floors. Lastly, around 68% of buildings are built before 1950.

The next most susceptible area to flood damage is 705.02, which is adjacent to 709.01. As a result, the area is very similar in terms of flood hazard and flood vulnerability to that of census tract 709.01. There are also many adjacent East Boston areas that are in the top 10 census tracts that may incur the most damage. All of these areas have high flood hazard risk due to low elevation and slope (since they are along or near a coast) as well as high land cover risk, which is also due to the fact that they are more urban areas near water. Additionally, it makes sense that areas that are adjacent to each other have similar levels of flood risk.

The areas that have low flood risk vulnerability are those that are in areas with a lot of greenspace and vegetation and limited concrete. For example, areas such as 9807, 1401.06, 1304.04 and 1303 have low flood risk due to being or being near highly vegetated areas such as the Stony Brook Reservation. Additionally, these census tracts have very high elevation values. Census tract 1303 has a normalized elevation value of 1 and census tract 1304.04 has a normalized elevation value of 0.84. Lastly, census tract 1401.06 has a normalized elevation value of 0.96.

4.5 Overall Flood Risk Model Evaluation

There have been very few works combining flood hazard features obtained from aerial images as well as vulnerability features from socioeconomic data and building datasets to find the rankings of susceptible urban areas as granular as census tracts. As a result, it is difficult to completely verify the accuracy of these rankings quantitatively. However, the city of Boston has created an interactive tool that has results from various flood models to predict the areas affected by different types of floods across

time frames in the future [1] [5].

Figure 4-6 displays flooding visualizations from the Boston Climate Ready Map for the top 10 census tracts that the Flood Risk model has output (shown in Table 4.12). The teal regions are the predicted long term storm water flooding areas. The light blue color represents areas impacted by coastal flooding as a result of a 36 inch rise in sea-level in the long term (2070s). For a majority of the top 10 high risk census tracts, the Boston Climate Ready models are in agreement. Many of these areas are at risk for flood hazard. The Boston Climate Ready map also shows visualizations of vulnerable populations such as the elderly, children, medical illness, disabilities, limited English proficiency, people of color, and low income. The visualizations (not included in this report) show more than one vulnerable population living in each of the high risk census tracts.

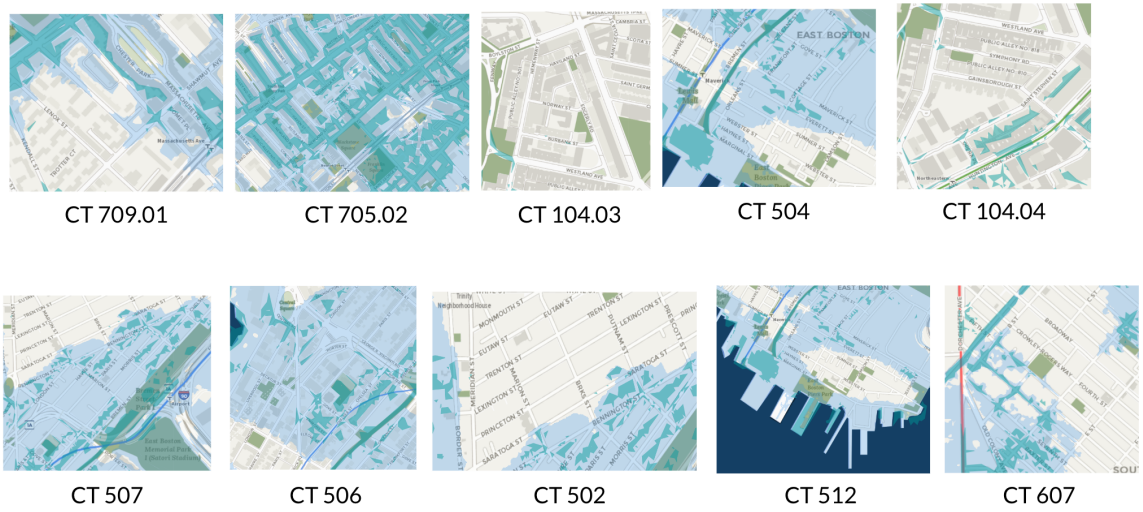

Figure 4-6: Boston Climate Ready Map Visualizations for each of the top 10 high flood risk (vulnerability + hazard) census tracts [5]

[1] https://www.boston.gov/departments/environment/climate-ready-boston-map-explorer

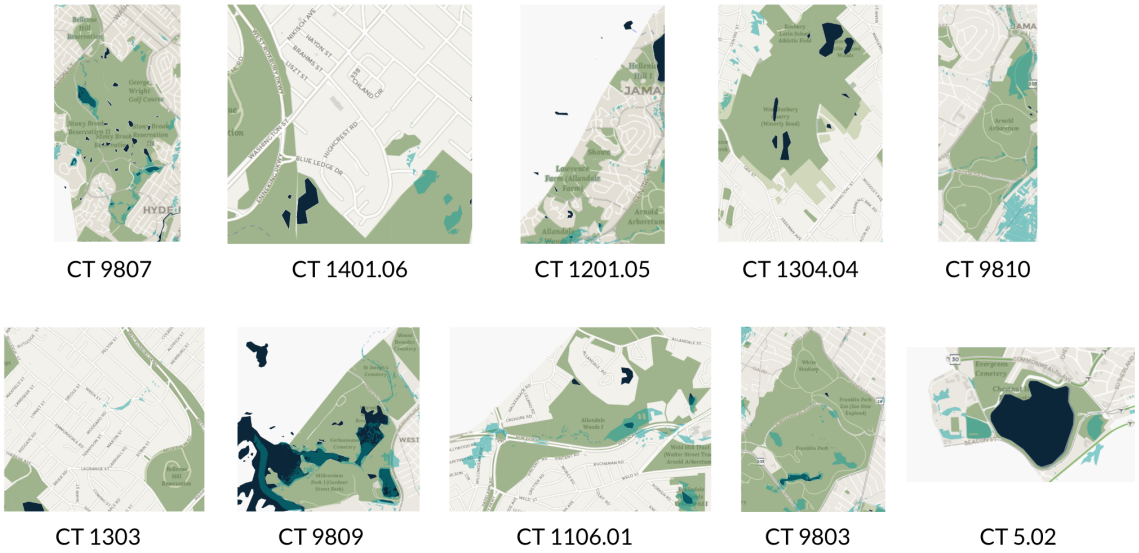

Figure 4-7: Boston Climate Ready Map Visualizations for each of the top 10 low flood risk (vulnerability + hazard) census tracts [5]

In contrast, Figure 4-7 shows visualizations of areas impacted by flooding for the 10 low-risk census tracts. These visualizations are also from the Boston Climate Ready Map. The color scheme is the same as in Figure 4-6, and the areas in green represent open-spaces in Boston. It is evident that there are a lot more open-spaces in the low risk areas in Figure 4-7 than in the Figure 4-6 images. Additionally, there are significantly smaller areas that are predicted to be affected by storm-water flooding, and virtually no areas that will be impacted by coastal flooding. Lastly, less of the low risk census tracts are covered by vulnerable population layers in the Boston Climate Ready Map. As a result, these areas are both low risk in terms of flood hazard and flood vulnerability. From these observations, we can see that the outputs from the flood risk model are qualitatively consistent with predictions made by the Boston Climate Ready map.

Chapter 5

Conclusion

5.1 Discussion

Climate change has been exacerbating flood frequency as well as intensity. As a result, many existing flood models that use historical data to create flood maps may not accurately capture the growing effects of climate change towards flooding. In this project we aggregated flood hazard and flood vulnerability features to build a granular flood risk model at the census tract level. Additionally, we considered the potential flood hazards from urban land cover by finetuning the Segmenter, KNet and PSPNet segmentation models on a dataset we curated consisting of images and segmentations from the Kaggle Aerial Drone Dataset, Nearmap, and Apple Maps. All three finetuned models yielded significantly improved results from the baseline. After evaluating the qualitative and quantitative results for all three model architectures, we found Segmenter to be the best performing model for the task. The finetuned Segmenter model had a mean IoU of 78% and had pixel accuracies of at least 70% for each class. Not only does the image segmentation model yield promising results for Boston images, but it is also able to generalize to areas in New York City. Many areas in New York City such as Manhattan are a lot more urban than Boston also include a higher density of taller buildings creating shadowed areas. However, the finetuned model handles these cases reasonably well.

We also manually collected georeferenced images from Nearmap and created a

masking algorithm to keep track of the pixels belonging to each census tracts in the images. Then, we applied the finetuned Segmenter model on the images to find areas of grass, trees, concrete/sidewalk, soil, and water in each census tract.

The features we aggregated for the model are also very relevant to flood hazard as well as vulnerability. For example, for flood vulnerability, we took into account both socioeconomic factors as well as building information. More specifically, we used datasets to find areas of socioeconomically vulnerable populations such as people who are low income, over the age of 65, people who are disabled, people who have limited English proficiency, people of color, children under the age of 6, and foreign born populations who are not US citizens.

We also utilized the Boston Buildings Dataset to incorporate building vulnerabilities in the model. For example, for each census tract in Boston, we found the number of buildings that are residential, below six floors tall, and built before 1950. Areas with a larger number of these features are more susceptible to flood risk. After combining the features with the relative weights and obtaining the highest and lowest risk census tracts, we found the corresponding areas on the Boston Climate Ready Map visualizer to qualitatively evaluate the high hazard and social vulnerability areas. We observe that a majority of the top 10 census tracts with the highest risk as per the flood risk model is shown to have a significant area of predicted flooded regions on the Boston Climate map. In contrast, the areas that have the lowest flood risk based on the flood risk model has little to no predicted coastal or storm-water flooded areas on the Boston Climate map.

5.2 Limitations

One of the limitations in our project is the relatively small dataset for finetuning. Although finetuning a pre-trained Segmenter model does not require a dataset as big as ADE20K (20,000 images), our dataset had a total of 545 images, where 432 of them were used for training. Another limitation in our project is that the finetuned model tends to classify a significant portion of images to be clutter. Often times, these are

areas occluded by shadows or that have low visibility. An improved segmentation model would be able to infer the patterns necessary to predict these low visibility areas. This may be possible by finetuning the model on a larger dataset containing more urban images.

The ordering of the weights for the land use were determined by those found in previous literature. However, it would be beneficial to find more precise weights using the analytical hierarchy process. Additionally, we currently weight all the socioeconomic values the same, but collaborating with an expert in the field may lead to more accurate weighting of these values.

5.3 Future Work

This project can be extended by incorporating more features related flood hazard and vulnerability risk. For example, although our image segmentation model is able to detect sidewalk areas, it would be beneficial to find the specific material of the sidewalk. This is because surface materials impact water runoff during flood events. More specifically, more permeable sidewalk materials are beneficial in the event of a flood. Finding sidewalk types can be done by using a semantic segmentation model called CitySurfaces [16]. Most cities do not have a catalogue of the different types of pavement surfaces and a computer vision approach is less time-consuming than the alternative manual process. The model is able to detect 8 kinds of sidewalk materials from street view images: concrete, bricks, asphalt, granite, mixed concrete & brick, hexagonal asphalt paver, and cobblestone roads. All of these sidewalk materials can be found in Boston and differentiating between them in our model would be beneficial.

Another flood related feature that would be important to use is soil type. Similar to sidewalk materials, some soil types are more permeable than others. Additionally, soil texture also plays a role in the amount of flooding. Areas that have soil with finer texture are more likely to be flooded than areas with coarser soil. Studies such as Negese et al. (2022) have split the following soil types as very low (Eutric Cambisols), low (Eutric Nitisols), moderate (Chromic Luvisols), high (Leptosols), and very high

(Chromic Vertisols and Pellic Vertisols) impact in flood susceptibility [21]. Using the information from previous studies like this can be useful to make the flood model more rich. Lastly, it would be beneficial to gain feedback and input from experts in the fields of flood analysis and other environmental domains.

Bibliography

[1] Aerial semantic segmentation drone dataset. 39

[2] Boston building inventory. 34, 35

[3] Boston climate vulnerability assessment. 15, 34, 35

[4] Census reporter. 63, 66, 67

[5] Climate ready boston map explorer. 12, 69, 70

[6] Climate ready boston social vulnerability. 36, 37

[7] The demographics of disaster. 36

[8] Intersect over union (iou). 11, 52

[9] Abul Kalam Azad, Khondoker Mokaddem Hossain, and Mahbuba Nasreen. Flood-induced vulnerabilities and problems encountered by women in northern bangladesh. *International Journal of Disaster Risk Science*, 4:190–199, 2013. 37

[10] Olivia Box. Could more urban trees mitigate runoff and flooding?, Aug 2021. 30

[11] MMSegmentation Contributors. MMSegmentation: Openmmlab semantic segmentation toolbox and benchmark. https://github.com/open-mmlab/mmsegmentation, 2020. 43

[12] Hamid Darabi, Ali Torabi Haghighi, Mohamad Ayob Mohamadi, Mostafa Rashidpour, Alan D. Ziegler, Ali Akbar Hekmatzadeh, and Bjørn Kløve. Urban flood risk mapping using data-driven geospatial techniques for a flood-prone case area in iran. *Hydrology Research*, 2019. 21

[13] Maggie Davis. Flood risk models vary widely - here's what you need to know (and how to mitigate threats), Jun 2022. 16

[14] Md. Nawrose Fatemi, Seth Asare Okyere, Stephen Kofi Diko, Michihiro Kita, Motoki Shimoda, and Shigeki Matsubara. Physical vulnerability and local responses to flood damage in peri-urban areas of dhaka, bangladesh. *Sustainability*, 2020. 35

[15] Queensland Government. What factors contribute to floods?, Jun 2018. 30

[16] Maryam Hosseini, Fábio Miranda, Jianzhe Lin, and Cláudio T. Silva. Citysurfaces: City-scale semantic segmentation of sidewalk materials. *ArXiv*, abs/2201.02260, 2022. 73

[17] Jian Kang, Marco Körner, Yuanyuan Wang, Hannes Taubenböck, and Xiaoxiang Zhu. Building instance classification using street view images. *ArXiv*, abs/1802.09026, 2018. 22

[18] Christopher P Konrad. Effects of urban development on floods, Nov 2016. 15, 16

[19] Björn Lütjens, Brandon Leshchinskiy, Christian Requena-Mesa, Farrukh Ahmed Chishtie, Natalia Díaz Rodríguez, Océane Boulais, Aruna Sankaranarayanan, Aaron Piña, Yarin Gal, Chedy Raïssi, Alexander Lavin, and Dava Newman. Physically-consistent generative adversarial networks for coastal flood visualization. *ArXiv*, abs/2104.04785, 2021. 21

[20] Shaun A. Maskrey, Nick J. Mount, and Colin R. Thorne. Doing flood risk modelling differently: Evaluating the potential for participatory techniques to broaden flood risk management decision-making. *Journal of Flood Risk Management*, 15, 2021. 21

[21] Ajanaw Negese, Dessale Worku, Alazar Shitaye, and Haile Getnet. Potential flood-prone area identification and mapping using gis-based multi-criteria decision-making and analytical hierarchy process in dega damot district, northwestern ethiopia. *Applied Water Science*, 12, 2022. 19, 23, 33, 74

[22] M. Panahi, Abolfazl Jaafari, Ataollah Shirzadi, Himan Shahabi, Omid Rahmati, Ebrahim Omidvar, Saro Lee, and Dieu Tien Bui. Deep learning neural networks for spatially explicit prediction of flash flood probability. *Geoscience frontiers*, 2020. 21, 22

[23] Mahfuzur Rahman, Chen Ningsheng, Md. Monirul Islam, Ashraf M. Dewan, Javed Iqbal, Rana Muhammad Ali Washakh, and Tian Shufeng. Flood susceptibility assessment in bangladesh using machine learning and multi-criteria decision analysis. *Earth Systems and Environment*, 3:585 – 601, 2019. 17

[24] Veerappan Ramesh and Sayed Sumaira Iqbal. Urban flood susceptibility zonation mapping using evidential belief function, frequency ratio and fuzzy gamma operator models in gis: a case study of greater mumbai, maharashtra, india. *Geocarto International*, 37:581 – 606, 2020. 31

[25] Albert Reuther, Jeremy Kepner, Chansup Byun, Siddharth Samsi, William Arcand, David Bestor, Bill Bergeron, Vijay Gadepally, Michael Houle, Matthew Hubbell, Michael Jones, Anna Klein, Lauren Milechin, Julia Mullen, Andrew Prout, Antonio Rosa, Charles Yee, and Peter Michaleas. Interactive supercomputing on 40,000 cores for machine learning and data analysis. In *2018 IEEE*

High Performance extreme Computing Conference (HPEC), pages 1–6. IEEE, 2018. 38

[26] Neil Martin Robertson and Tak Chan. Aerial image segmentation for flood risk analysis. *2009 16th IEEE International Conference on Image Processing (ICIP)*, pages 597–600, 2009. 25

[27] Cassie Roopnarine, Bheshem Ramlal, and Ronald Roopnarine. A comparative analysis of weighting methods in geospatial flood risk assessment: A trinidad case study. *Land*, 2022. 17, 31

[28] Robin Strudel, Ricardo Garcia Pinel, Ivan Laptev, and Cordelia Schmid. Segmenter: Transformer for semantic segmentation. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7242–7252, 2021. 11, 25, 44

[29] Matej Vojtek and Jana Vojteková. Flood susceptibility mapping on a national scale in slovakia using the analytical hierarchy process. *Water*, 2019. 23, 32, 33

[30] Qi Wen, Kaiyu Jiang, Weiqi Wang, Qingjie Liu, Qing Guo, Lingling Li, and Ping Wang. Automatic building extraction from google earth images under complex backgrounds based on deep instance segmentation network. *Sensors (Basel, Switzerland)*, 19, 2019. 22

[31] Wenwei Zhang, Jiangmiao Pang, Kai Chen, and Chen Change Loy. K-net: Towards unified image segmentation. In *NeurIPS*, 2021. 11, 24, 44

[32] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6230–6239, 2017. 24

[33] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. Scene parsing through ade20k dataset. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5122–5130, 2017. 11, 38, 45