이학석사학위논문

Boundary-Driven Adversarial Learning of Deep Neural Networks for Flood Monitoring using SAR

SAR 영상을 활용한 홍수 모니터링을 위한 경계 강화 적대적 학습 딥러닝 모델

2023년 8월

서울대학교 대학원

지구환경과학부

김 휘 송

Boundary-Driven Adversarial Learning of Deep Neural Networks for Flood Monitoring using SAR

SAR 영상을 활용한 홍수 모니터링을 위한
경계 강화 적대적 학습 딥러닝 모델

지도 교수 김 덕 진

이 논문을 이학석사 학위논문으로 제출함
2023 년 5 월

서울대학교 대학원
자연과학대학 지구환경과학부
김 휘 송

김휘송의 이학석사 학위논문을 인준함
2023 년 7 월

위 원 장 ______ 조 양 기 ______ (인)

부위원장 ______ 김 덕 진 ______ (인)

위 원 ______ 박 상 은 ______ (인)

Abstract

The necessity of real-time flood monitoring has been increasing as the frequency and intensity of water-related disasters increase. Synthetic Aperture Radar(SAR) could be particularly useful for a inundation mapping because it is able to penetrate clouds and provide images even during periods of darkness. Therefore, water segmentation using SAR has been actively researched, especially the advent of Convolutional Neural Networks(CNN) contributing to high overall accuracy. However, CNN is vulnerable to detecting precise boundaries and narrow rivers, which pose challenges for practical applications. In this study, we propose a boundary-driven adversarial learning approach of deep neural networks to detect waterbodies with precise borders and small rivers. We adopt the adversarial learning of Generative Adversarial Networks (GAN) to make a generator focus on pixels that could be easily ignored. A discriminator evaluates and distinguishes the segmented images with the ground truth labels by consulting the SAR images and the boundary distance map. The boundary distance map is designed to highlight the small area like the boundaries and streams and suppressing false positive errors.Moreover, we propose a hybrid loss function that guides the network to concentrate on both the overall and the fine details by fusing Binary Cross Entropy loss, Hausdorff distance loss and adversarial loss. Through adversarial training with the hybrid loss, the water segmentation model using SAR can precisely detect waterbodies. We demonstrate the effectiveness of the model using three evaluation metrics: F1-score, Boundary IoU, and Matthews Correlation Coefficient, and we also apply additional qualitative assessment. Our empirical evidence indicates that the proposed model outperforms other segmentation models like U-Net and

DeepLabv3+, especially in terms of precise boundaries and narrow rivers. To assess the practical monitoring use, we demonstrate that the proposed model maintains precision with the large scene SAR images. Not only does it detect precise boundaries and narrow objects,but it also reduces false positive errors in large scene SAR images. The visual inspection further demonstrates that our model can detect narrow rivers and small reservoirs that are missed by other models, showcasing the potential of boundary-driven adversarial learning of deep neural networks in practical monitoring use.

Keywords: Water Segmentation, SAR, Adversarial Learning, Flood Monitoring

Student ID: 2021-29595

Contents

Abstract **i**

Contents **iii**

List of Figures **v**

List of Tables **vi**

1 Introduction **1**

1.1 Research Background 1

1.2 Purpose of Research 7

2 Data Acquisition **11**

3 Boundary-driven Adversarial Learning of Deep Neural Networks **21**

3.1 Generator architecture 23

3.2 Discriminator architecture 25

3.3 Hybrid Loss 30

4 Experiments **33**

4.1 Experiment Settings 33

4.2 Evaluation Metrics 36

4.3 Comparison to other segmentation models 38

4.4 Ablation Studies 43

5 Discussion **51**

6 Conclusion **57**

Bibliography **59**

초 록 **64**

List of Figures

1.1 Sample results of our proposed method compared to U-Net and DeepLabV3+. 10

2.1 The procedure of constructing training dataset from Landcover Map. 15

2.2 Study area and data coverage near Korea and Southeast Asia. . . 19

2.3 The examples of the final dataset (a) Lnadcover map based dataset (b) UNOSAT flood map based dataset. 20

3.1 The full flowchart of the SAR water segmentation with adversarial learning. 22

3.2 The detailed architecture of the SAR water segmentation generator.24

3.3 The detailed architecture of the SAR water segmentation discriminator. 27

3.4 The procedure of constructing the boundary distance map. 28

3.5 The examples of the boundary distance map with groundtruth label images. 29

4.1 Illustration of F1-score and MCC on three models: U-Net, Deeplabv3+ and the proposed model. 40

4.2 Qualitative performance evaluation of the SAR Water Segmentation. 42

4.3 Illustration of different reconstruction loss functions(L_{Recon}) and corresponding bound loss(L_{Bound}) and total segmentation loss(L_{Seg}) for each training epoch. 48

5.1 Application results in Korea Peninsula of the waterbodies monitoring. 55

5.2 The results of the three models by the incidence angle difference 56

List of Tables

2.1 List of Sentinel-1 SAR satellite images for training data procurement 16

4.1 Implementation details of experiments 35

4.2 Performance comparison of SAR Water Segmentation 41

4.3 Ablation study on discriminator and constraints 44

4.4 Ablation study on different reconstruction loss functions and the loss rescaling parameters 47

4.5 Ablation study of the Model Architecture on different bound loss rescaling parameter α 50

5.1 The detailed information of Sentinel-1 SAR images for application test in Korean Peninsula 54

1 Introduction

1.1 Research Background

As the climate change accelerates, the frequency and intensity of extreme weather events including floods and inundations are likely to increase. As the unpredictability of the weather increases, the significance of the response and recovery in the disaster response system is accentuated compared to the previous state. For the rapid response and recovery, the real-time and precise inundation monitoring is in demand. As satellites periodically provides the remote sensing data with large-scale coverage, flood monitoring based on the remote sensing could help track changes in the frequency and intensity of floods and inundations in different regions over time. Additionally, it can evaluate the effectiveness of flood management measures. As part of the fundamental research for flood monitoring, studies on waterbody detection are actively progressing using multispectral sensor, Light Detection And Ranging(LiDAR), and Synthetic Aperture Radar(SAR)(Dong et al. 2019; Höfle et al. 2009; White et al. 2015; Yuan et al. 2021).

SAR is widely adopted for detecting waterbodies, such as lakes, coastlines and rivers. SAR, an active microwave sensor, can penetrate clouds and does not require sunlight, so SAR images can be obtained regardless of weather and time. This advantage enables SAR as a fundamental data for water monitoring(Pradhan et al. 2017). SAR emits the microwaves and records the reflected microwaves from the surface, so the roughness of the surface has a significant effect on the amplitude of SAR images. Specifically, the water surface is usually smooth, causing the microwaves to specularly reflect away with few back scatterings. On the other hand, the land, which is typically rough, reflects the microwave back to the SAR sensor. This difference makes the waterbodies easily distinguish in SAR images. As a result, research on waterbody segmentation using SAR images has been rapidly expanding.

Traditional Machine Learning algorithms has been introduced in the field of waterbody segmentation methods using SAR images. SAR water segmentation studies based on Edge detection(Liu and Jezek 2004), Level Set method(Silveira and Heleno 2008), Clustering(Liu et al. 2016), Random Forest(Xie et al. 2015), and Support Vector Machine(Klemenjak et al. 2012) have obtained promising performances. However, the distinctive imaging processing of SAR leads to the presence of speckle noises. Moreover, the amplitude of pixels in one SAR image varies with the incidence angle. Traditional Machine Learning algorithms aim to define a formula for categorizing the pixels, which is easily disrupted by the speckle noises and varying amplitudes. Furthermore, from the perspective of the practical flood monitoring, Traditional Machine Learning algorithms usually require large computing power and memory resources, leading to slow analysis for real-time monitoring(Guo et al. 2022).

Fortunately, due to its powerful feature extraction capabilities, Deep Learning-based Convolutional Neural Networks(CNN) have shown remarkable results in SAR waterbody segmentation. As CNN extracts shared weights by iterating through all pixels, it can construct a formula that fits all pixels in SAR images regardless of speckle noises. As SAR images intrinsically contain a large amount of speckle noises, segmentation with whole SAR images is not easy(Guo et al. 2022). However, with the advent of GPUs and the Big Data, Deep Learning methods have been able to overcome this limitation by learning nonlinear relationships. CNN, which iterates the shared weights through multi-layer structures, is able to find robust weights that are unaffected by speckle noises. (Long et al. 2015) proposed Fully Convolutional Networks(FCN) by replacing the full connection layer in the tail of the model with the fully convolutional layer, enabling pixel-level segmentation while preserving the resolution. (Kang et al. 2018) firstly introduced FCN in the water body segmentation area, verifying that FCN is less sensitive to speckle noise of SAR images. Subsequently, U-Net(Ronneberger et al. 2015), which is composed of a U-shaped symmetrical structure with skip-connections, has become a popular deep learning model in the SAR waterbody segmentation (Denbina et al. 2020; Kim et al. 2021; Pai et al. 2019). U-Net extracts the deep and specific details of the waterbody in the encoder and inclusive and semantic characteristics in the decoder and connects this information through the skip-connections, which enhances the accuracy while maintaining the resolution of input images. In 2018, as the lastest version of Deeplab series, (Chen et al. 2018) proposed the DeepLabv3+ model to address the limitations in handling boundary details. Through implementing Atrous Spatial Pyramid Pooling(ASPP) and depthwise separable convolution with encoder-decoder structure, DeepLabv3+ can capture the multi-scale context and the sharper object boundaries.

However, these conventional CNN models tend to aggregate information at the expense of some object details, leading to inadequate segmentation of fine and narrow waterbodies(Guo et al. 2022). There are two ways to address this problem. Modifying the model architecture is applicable to universal regions, but it is more complex and time-consuming in practical use. On the other hand, adding auxiliary data directly to input data can indicate key areas of focus with intuitively interpretable information. However, this method struggles in constructing auxiliary data with consistent qualities in practical use for universal areas. For practical use in flood monitoring, the convenient construction of data with short prediction time is needed.

Meanwhile, (Goodfellow et al. 2020) proposed the concept of adversarial training in image generation tasks. Generative Adversarial Networks(GANs) have an additional network named the discriminator, which distinguishes whether the generated image is within the real data distribution or not(Goodfellow et al. 2020). On the contrary, the generative network tries to fool the discriminative network by generating real-like, exquisite images. As the two networks compete with each other in an adversarial setting, the generator learns based on the loss function derived from the discriminator's output. As the training progresses, the generator's output becomes more indistinguishable from the real data. Instead of inputting a random noise into regular GANs, the Conditional GANs(cGANs) receives additional data that conditions the models(Mirza and Osindero 2014). This condition can be a multi-modal, such as images, which (Isola et al. 2017) presents the image-image translation based on the U-Net generator. Adding of a reconstruction loss like L1 loss to the adversarial loss makes the generated image realistic, as demonstrated by (Isola et al. 2017).

The usage of GAN is not only restricted to image generation but also extends to the field of image segmentation. In the task of the segmentation, the combination of GANs and CNN results in a synergy effect, generating more realistic and detailed segmentation maps. This hybrid approach has been actively researched in the field of the medical image segmentation, where input images are challenging to analyze and the details of segmentation is critical. (Lei et al. 2020) constructed a skin lesion segmentation model based on dual discriminators with different perspectives to augment precise decision. (Park et al. 2019) proposed an endometrium segmentation model trained with the key-point discriminator, which surmounts the challenges of ambiguous boundaries and heterogeneous textures in transvaginal ultrasound images. Based on the ground truth endometrium key-point map and the intermediate product of the generator, the key-point discriminator determines whether the segmentation results coincide with the input ultrasound images.

The application of adversarial learning in SAR segmentation task is still in its early stages, with only few studies. In 2020, (Ronci et al. 2020) introduced the adversarial learning for oil spill detection using Radarsat-2 and Sentinel-1 SAR images. They compared with standard U-Net and U-Net++(Zhou et al. 2018), and the U-Net with adversarial training showed the best accuracy. (Li et al. 2021) overcame the limitation of a small oil spill train dataset by constructing a series of adversarial networks at multiple scales. The proposed Multiscale Conditional Adversarial Network(MCAN) outperformed adaptive thresholding, level set, CGAN, FCN, and U-Net. (Li et al. 2022) utilized the cycle-consistent adversarial network(CycGAN)(Zhu et al. 2017) to generated optical images from SAR images, and two images were stitched together to segment marine culture farms. With the attention module(Vaswani et al. 2017) in segmentation, the model precisely extracts regularly shaped objects. (Fan and Liu 2023) designed a MultiTask Generative Adversarial Networks(MTGANs) to discriminate oil spills and look-alikes and extract oil spill areas. The first GAN model is trained to generate SAR oil spill images, and the discriminator of this model distinguishes real oil and look-alike images. Another GAN model segments the oil spill area, and SAR images, generated segmentations and segmentation labels are fed into the discriminator to distinguish real from generated labels. MTGANs is superior compared to other state-of-the-arts models, especially in maintaining details and reducing misclassification.

To the best of our knowledge, no previous study has proposed a method of adversarial learning in SAR water segmentation. Adversarial learning using SAR images is currently only adopted in the area of SAR-to-optical or optical-to-SAR image translation. Therefore, we sought inspiration from the field of medical image segmentation, where segmentation is based on non-optical images and the boundary is critical information.

1.2 Purpose of Research

As mentioned earlier, the performance of vanilla CNNs in precise details is not satisfactory(Guo et al. 2022). There are two main challenges in accurate Deep-Learning based SAR water segmentation: (1) The area of low intensity is easily misjudged as water. As the SAR water segmentation is based on the difference of brightness in SAR images, the other dark regions which are flat surfaces(e.g. golf courses, road) or the radar shadow in complex topography are also misclassified as water(Figure 1.1. (a)). (2) The narrow rivers are easily ignored on the basis of the SAR resolution and the detailed structures of water bodies are insufficient(Figure 1.1. (b)). This is because the receptive field of the convolutional layers in CNN gradually increases with deeper layers, which can result in a loss of fine spatial details and subtle boundaries.

To overcome the above discrepancy, we introduce boundary-driven adversarial learning to the SAR water segmentation model. We propose a novel Deep Learning framework that comprises a segmentation network and a boundary-guided discriminator network. We adopt the U-Net architecture as the main segmentation model and substitute pooling layers with strided convolutional layers for a simpler but deeper learning. Additionally, we introduce a novel hybrid loss that fuses Binary Cross Entropy loss, Hausdorff distance loss, and adversarial loss to address the intrinsic weaknesses of region-based loss functions. The proposed discriminator is guided by the SAR images and the boundary distance map of ground truth images. The boundary-driven discriminator determines whether the predicted segmentation results, based on the ground truth boundary distance map, align with the actual boundaries. The ground truth boundary map is the normalized Euclidean distance from the thickened waterbody contour. By sequentially thickening and calculating distance, contours of narrow rivers are merged and highlighted, while contours of wider rivers do not intersect. Through adversarial learning, the proposed model strives to predict the waterbody more accurately, especially along borders and narrow rivers.

The main contributions of this article are given as follows: 1) A boundary-aware adversarial learning method is proposed for SAR Water Segmentation. We introduce the discriminator into the SAR Water Segmentation task, allowing model to learn a detailed formula with ancillary data during training without the need for inference. The adversarial learning approach mitigates the difficulties of constructing consistent ancillary data for wide areas and reduces the time-consuming data processing required for real-time monitoring. 2) The proposed SAR water segmentation model, with the modified architecture and hybrid loss, is guided by the waterbody boundary distance map. This enables us to extract the waterbody with higher accuracy, especially along coincident borders and in detecting narrow rivers.

The remainder of the dissertation introduces the boundary-driven adversarial learning-based SAR water segmentation model. Chapter 2 provides the procedure for constructing the training dataset, which includes Sentinel-1 SAR image, Landcover Map and UNOSAT Disaster Flood Map. Chapter 3 presents the details of our proposed model, including the SAR water segementation model architecture, discriminator architecture and hybrid loss. The experiment settings and evaluation metrics are discussed in Chapter 4. Furthermore, Chapter 4 reports the qualitative and quantitative results of proposed boundary-driven adversarial learning based SAR water segmentation model, comparing it to other segmentation models. To identify how different parts of the model interact and contribute, the results of ablation studies are also presented in Chapter 4. Chapter 5 is dedicated to the discussion of the results, and the final chapter concludes the dissertation, summarizing the key points and contributions of the studies.

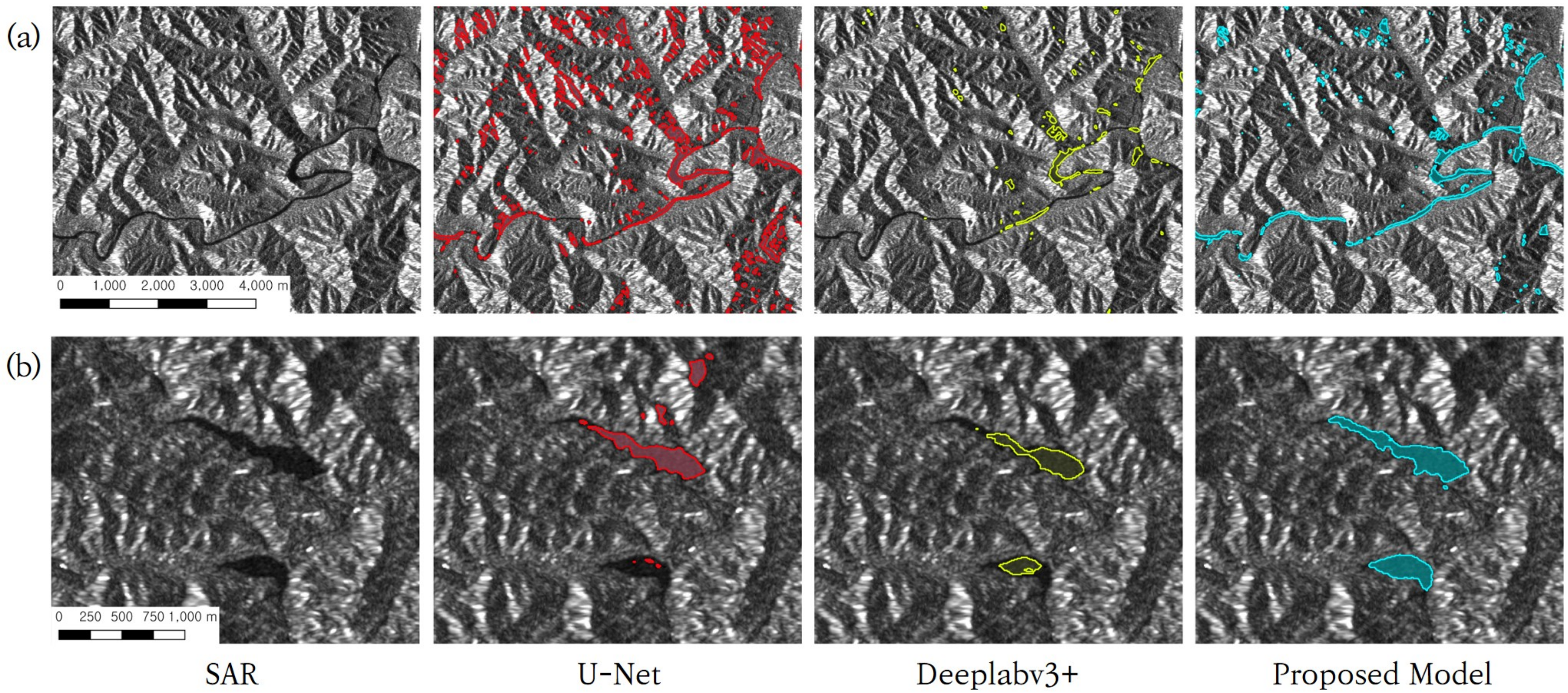

Figure 1.1 Sample results of our proposed method compared to U-Net and DeepLabV3+. (a)the region of small reservoir (b)the region of small stream surrounded by moundtain. From left to right, SAR images, and the result of U-Net, DeepLabV3+ and the proposed model are displayed.

2 Data Acquisition

As the purpose of this research is to create a high-quality water segmentation Deep Learning model using SAR images, it is necessary to construct the training and test datasets. In this study, Sentinel-1 SAR images and water ground truth data, using Landcover Map and Flood Map, were adopted to generate the dataset. This chapter explains the series of procedures to construct the dataset. Figure 2.1. describes the whole procedure of the constructing dataset.

Sentinel-1 satellite was selected for SAR data due to its high accessibility. Utilizing remote sensing data for flood monitoring can incur high costs, making it one of the biggest obstacles. In this regard, Sentinel-1, which is freely offered by European Space Agency(ESA), is the optimal SAR satellite. Sentinel-1 carries a C-band SAR instrument and usually operates in Interferometric Wide swath(IW) mode over land and coastal areas. We acquired the SAR images in IW mode Level-1 High Resolution Ground Range Detected(GRD-H) which has a resolution of 20m x 22m.

The dataset is composed of input SAR images and corresponding water-labeled images. The difficulty in constructing the dataset lies in preparing a high-quality water surface map. To overcome this difficulty, we utilized the Landcover Map from the Ministry of Environment of South Korea and the Disaster Flood Map from the United Nations Satellite Centre(UNOSAT). As the formation of two data is different, the procedure to obtain the corresponding Sentinel-1 was also different.

For the Landcover Map, we select the subdivision Landcover Map, which has a 1m resolution and is categorized into 41 landcover types. The Ministry of Environment produces the Landcover Map based on aerial orthogonal images along with other ancillary data such as Digital Topographic Map, Cadastral Map, and Forest Cover Map. The Landcover Map is constructed by on-screen digitizing by visual inspection. Therefore, the Sentinel-1 images that coincide with the date and region of the referred aerial orthogonal images can be utilized as the corresponding SAR images for groundtruth data from the Landcover Map. Taking inspiration from this idea, we accumulated the shooting date of aerial images for each map grid through visual examination and data crawling. Since the recycle date of Sentinel-1 is 12 days, obtaining exact matching images is challenging. Hence, we lowered the criteria to plus or minus three days from the aerial shooting date.

For the Flood Map, UNOSAT's Emergency Mapping service provides satellite-based disaster analysis of floods, landslides, earthquakes, volcanoes and other disasters. Satellite data is not only restricted to Sentinel-1, but also includes other SAR and optical satellites such as VIIRS, RADARSAT Constellation Mission, and Gaofen-3. However, we only collected the Flood Map that analyzed the Sentinel-1 satellite images. Yet, the reliability of the Flood Map is unproven, so we could not utilize those maps directly. Therefore, we sorted out the high-quality Flood Map through visual inspection. By obtaining the the date and extent of the region from the Flood Map, we could identify the specific Sentinel-1 image that was analyzed.

Table 2.1 shows the list of Sentinel-1 images utilized for constructing the training dataset, and Figure 2.2 describes the corresponding regions for each Sentinel-1 image. Once the list of Sentinel-1 images was prepared, we downloaded them from Alaska Satellite Facility(ASF) Vertex. The IW mode of Sentinel-1 supports both single polarization(HH or VV) and dual polarization(HH+HV, VV+VH), but we select dual polarization for pretest and training. By previous empirical data, VV polarization was found to be the best for detecting water.

For preprocessing of Sentinel-1 GRDH images, we utilized Sentinel Applications Platform(SNAP) program, which is offered by ESA. We sequentially remove GRD-border noise and then apply radiometric calibration. Radiometric calibration is a necessary step that allows the measurement of radar backscatter to be converted into meaningful physical units, such as power or amplitude. Lastly, we apply terrain correction into a 10m resolution using SRTM 1sec HGT DEM based on EPSG 4326 projection.

We matched the preprocessed Sentinel-1 images and ground truth label shapefiles by date and regions. We cropped the preprocessed Sentinel-1 images that overlapped with label shapefiles. For the ground truth shapefiles, we rasterized them into the same projection and resolution as the corresponding Sentinel-1 images. After matching, the dataset contains few meaningless training sets where SAR images were at the border of the whole Sentinel-1 images. We eliminated them and also excluded pairs that did not cincide with each other through visual examination. The final dataset was cropped into 256 * 256 pixels. After the whole process was completed, the final dataset consisted of 28,569 pairs of the water ground truth images and SAR images. Figure 2.3

shows the 50 examples of the final dataset for landcover-based and flood-based dataset, respectively.

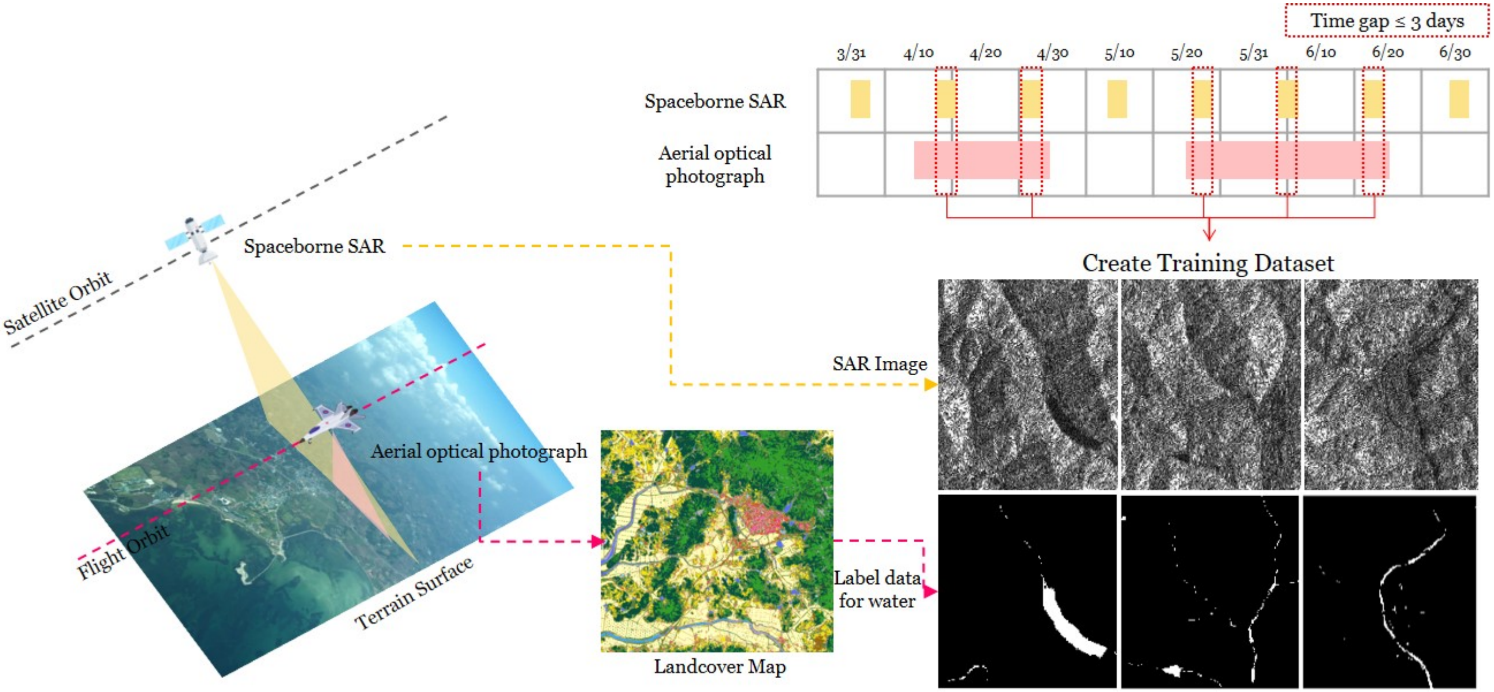

Figure 2.1 The procedure of constructing training dataset from Landcover Map.

Table 2.1 List of Sentinel-1 SAR satellite images for training data procurement

SAT	Mode, Type	Acquisition Time(UTC)	Orbit Dir	SAT	Mode, Type	Acquisition Time(UTC)	Orbit Dir
S1A[1]	IW, GRDH	25 Apr 2017 9:30:54 - 9:31:23	ASC[2]	S1A	IW, GRDH	30 Jul 2017 11:04:20 - 11:04:45	ASC
S1A	IW, GRDH	2 May 2017 9:22:51 - 9:23:21	ASC	S1A	IW, GRDH	15 Jun 2018 23:47:18 - 23:47:43	DESC[3]
S1A	IW, GRDH	7 May 2017 9:30:54 - 9:31:24	ASC	S1A	IW, GRDH	25 Jul 2018 11:04:26 - 11:04:51	ASC
S1A	IW, GRDH	14 May 2017 9:22:51 - 9:23:20	ASC	S1A	IW, GRDH	11 Jul 2015 11:54:34 - 11:54:59	ASC
S1A	IW, GRDH	19 May 2017 9:30:55 - 9:31:24	ASC	S1A	IW, GRDH	6 Aug 2015 11:37:30 - 11:37:55	ASC
S1A	IW, GRDH	7 Jun 2017 9:22:53 - 9:23:22	ASC	S1A	IW, GRDH	6 Aug 2015 11:37:55 - 11:38:20	ASC
S1A	IW, GRDH	12 Jun 2017 9:30:56 - 9:31:25	ASC	S1A	IW, GRDH	11 Aug 2015 11:47:21 - 11:47:46	ASC
S1A	IW, GRDH	19 Jun 2017 9:22:53 - 9:23:23	ASC	S1A	IW, GRDH	30 Jun 2016 23:55:28 - 23:55:53	DESC
S1A	IW, GRDH	30 Aug 2017 9:22:57 - 9:23:27	ASC	S1A	IW, GRDH	24 Jul 2016 23:55:29 - 23:55:54	DESC
S1A	IW, GRDH	4 Sep 2017 9:31:01 - 9:31:30	ASC	S1A	IW, GRDH	7 Nov 2017 22:45:31 - 22:45:56	DESC
S1A	IW, GRDH	16 Sep 2017 9:31:01 - 9:31:30	ASC	S1A	IW, GRDH	18 Jul 2015 11:47:20 - 11:47:45	ASC
S1A	IW, GRDH	29 Oct 2017 9:22:59 - 9:23:28	ASC	S1B[4]	IW, GRDH	19 Apr 2017 21:31:10 - 21:31:35	DESC
S1A	IW, GRDH	14 Jun 2018 9:23:29 - 9:23:57	ASC	S1B	IW, GRDH	26 Apr 2017 21:23:14 - 21:23:39	DESC
S1A	IW, GRDH	19 Jun 2018 9:31:32 - 9:31:57	ASC	S1B	IW, GRDH	26 Apr 2017 21:23:39 - 21:24:12	DESC
S1A	IW, GRDH	1 Jul 2018 9:31:33 - 9:31:58	ASC	S1B	IW, GRDH	1 May 2017 21:31:11 - 21:31:36	DESC
S1A	IW, GRDH	25 Jul 2018 9:31:34 - 9:31:59	ASC	S1B	IW, GRDH	1 May 2017 21:31:36 - 21:32:01	DESC
S1A	IW, GRDH	6 Aug 2018 9:31:06 - 9:31:35	ASC	S1B	IW, GRDH	1 May 2017 21:32:01 - 21:32:26	DESC
S1A	IW, GRDH	6 Aug 2018 9:31:35 - 9:32:00	ASC	S1B	IW, GRDH	20 May 2017 21:23:15 - 21:23:40	DESC

[1] Sentinel-1A
[2] Ascending orbit direction
[3] Descending orbit direction
[4] Sentinel-1B

S1A	IW, GRDH	18 Aug 2018 9:31:07 - 9:31:36	ASC	S1B	IW, GRDH	20 May 2017 21:23:40 - 21:24:13	DESC
S1A	IW, GRDH	11 Sep 2018 9:31:08 - 9:31:37	ASC	S1B	IW, GRDH	25 May 2017 21:31:37 - 21:32:02	DESC
S1A	IW, GRDH	23 Sep 2018 9:31:08 - 9:31:37	ASC	S1B	IW, GRDH	25 May 2017 21:32:02 - 21:32:27	DESC
S1A	IW, GRDH	23 Sep 2018 9:31:37 - 9:32:02	ASC	S1B	IW, GRDH	1 Jun 2017 21:23:16 - 21:23:41	DESC
S1A	IW, GRDH	3 Apr 2019 9:31:06 - 9:31:35	ASC	S1B	IW, GRDH	6 Jun 2017 21:31:38 - 21:32:03	DESC
S1A	IW, GRDH	10 Apr 2019 9:23:02 - 9:23:31	ASC	S1B	IW, GRDH	13 Jun 2017 21:23:16 - 21:23:41	DESC
S1A	IW, GRDH	15 Apr 2019 9:31:06 - 9:31:35	ASC	S1B	IW, GRDH	18 Jun 2017 21:31:13 - 21:31:38	DESC
S1A	IW, GRDH	27 Apr 2019 9:31:07 - 9:31:36	ASC	S1B	IW, GRDH	18 Jun 2017 21:31:38 - 21:32:03	DESC
S1A	IW, GRDH	4 May 2019 9:23:03 - 9:23:32	ASC	S1B	IW, GRDH	5 Aug 2017 21:31:41 - 21:32:06	DESC
S1A	IW, GRDH	9 May 2019 9:31:07 - 9:31:36	ASC	S1B	IW, GRDH	24 Aug 2017 21:23:20 - 21:23:45	DESC
S1A	IW, GRDH	16 May 2019 9:23:04 - 9:23:33	ASC	S1B	IW, GRDH	29 Aug 2017 21:31:42 - 21:32:07	DESC
S1A	IW, GRDH	21 May 2019 9:31:08 - 9:31:37	ASC	S1B	IW, GRDH	22 Sep 2017 21:31:43 - 21:32:08	DESC
S1A	IW, GRDH	21 May 2019 9:31:37 - 9:32:02	ASC	S1B	IW, GRDH	29 Sep 2017 21:23:22 - 21:23:47	DESC
S1A	IW, GRDH	28 May 2019 9:23:04 - 9:23:33	ASC	S1B	IW, GRDH	11 Oct 2017 21:23:47 - 21:24:20	DESC
S1A	IW, GRDH	28 May 2019 9:23:33 - 9:23:45	ASC	S1B	IW, GRDH	16 Oct 2017 21:31:44 - 21:32:09	DESC
S1A	IW, GRDH	2 Jun 2019 9:31:08 - 9:31:37	ASC	S1B	IW, GRDH	23 Oct 2017 21:23:47 - 21:24:20	DESC
S1A	IW, GRDH	2 Jun 2019 9:31:37 - 9:32:02	ASC	S1B	IW, GRDH	15 Dec 2017 21:32:08 - 21:32:33	DESC
S1A	IW, GRDH	14 Jun 2019 9:31:09 - 9:31:38	ASC	S1B	IW, GRDH	14 Apr 2018 21:31:16 - 21:31:41	DESC
S1A	IW, GRDH	8 Jul 2019 9:31:10 - 9:31:39	ASC	S1B	IW, GRDH	14 Apr 2018 21:31:41 - 21:32:06	DESC
S1A	IW, GRDH	1 Aug 2019 9:31:12 - 9:31:41	ASC	S1B	IW, GRDH	26 Apr 2018 21:31:17 - 21:31:42	DESC
S1A	IW, GRDH	20 Aug 2019 9:23:10 - 9:23:39	ASC	S1B	IW, GRDH	26 Apr 2018 21:31:42 - 21:32:07	DESC
S1A	IW, GRDH	1 Sep 2019 9:23:20 - 9:23:45	ASC	S1B	IW, GRDH	26 Apr 2018 21:32:07 - 21:32:32	DESC
S1A	IW, GRDH	13 Sep 2019 9:22:56 - 9:23:21	ASC	S1B	IW, GRDH	8 May 2018 21:31:42 - 21:32:07	DESC
S1A	IW, GRDH	18 Sep 2019 9:31:14 - 9:31:43	ASC	S1B	IW, GRDH	20 May 2018 21:31:43 - 21:32:08	DESC

S1A	IW, GRDH	25 Sep 2019 9:23:40 - 9:24:05	ASC	S1B	IW, GRDH	20 May 2018 21:32:08 - 21:32:33	DESC
S1A	IW, GRDH	7 Oct 2019 9:23:11 - 9:23:40	ASC	S1B	IW, GRDH	1 Jun 2018 21:31:19 - 21:31:44	DESC
S1A	IW, GRDH	7 Oct 2019 9:23:40 - 9:24:05	ASC	S1B	IW, GRDH	1 Jun 2018 21:31:44 - 21:32:09	DESC
S1A	IW, GRDH	12 Oct 2019 9:31:15 - 9:31:44	ASC	S1B	IW, GRDH	1 Jun 2018 21:32:09 - 21:32:34	DESC
S1A	IW, GRDH	24 Oct 2019 9:31:15 - 9:31:44	ASC	S1B	IW, GRDH	14 Aug 2019 21:23:58 - 21:24:31	DESC
S1A	IW, GRDH	5 Nov 2019 9:31:15 - 9:31:44	ASC	S1B	IW, GRDH	13 Oct 2019 21:24:00 - 21:24:33	DESC
S1A	IW, GRDH	5 Nov 2019 9:32:09 - 9:32:34	ASC	S1B	IW, GRDH	25 Oct 2019 21:23:35 - 21:24:00	DESC
S1A	IW, GRDH	17 Nov 2019 9:32:09 - 9:32:34	ASC	S1B	IW, GRDH	6 Nov 2019 21:23:10 - 21:23:35	DESC
				S1B	IW, GRDH	6 Nov 2019 21:23:35 - 21:24:00	DESC

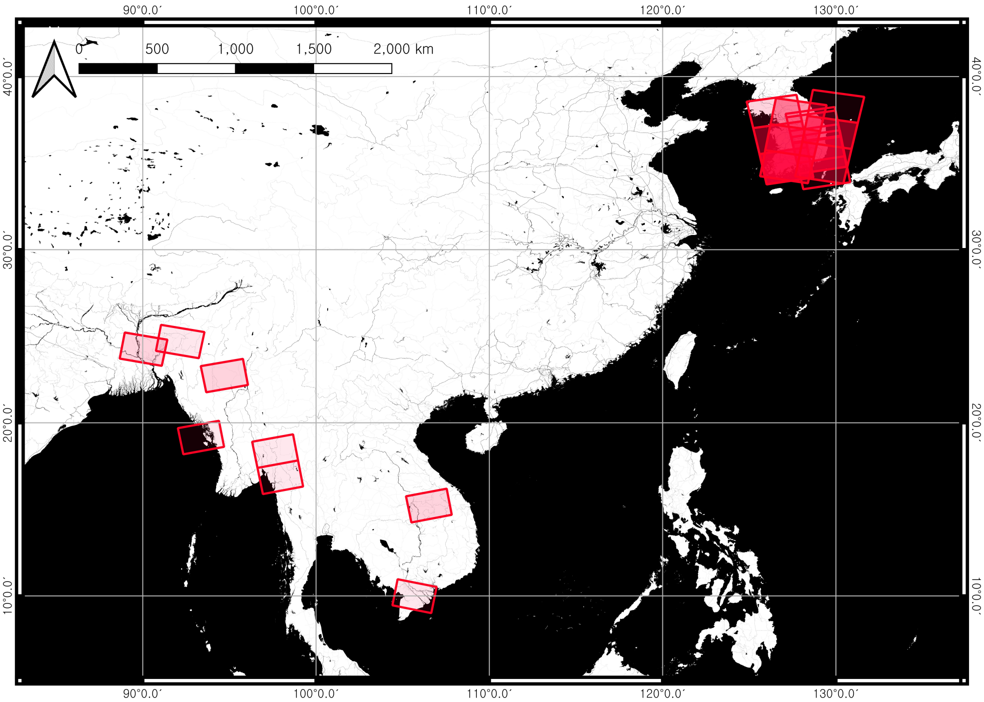

Figure 2.2 Study area and data coverage near Korea and Southeast Asia. Red boxes describe the spatial coverage of Sentinel-1 images including both ascending and descending paths.

(a)

(b)

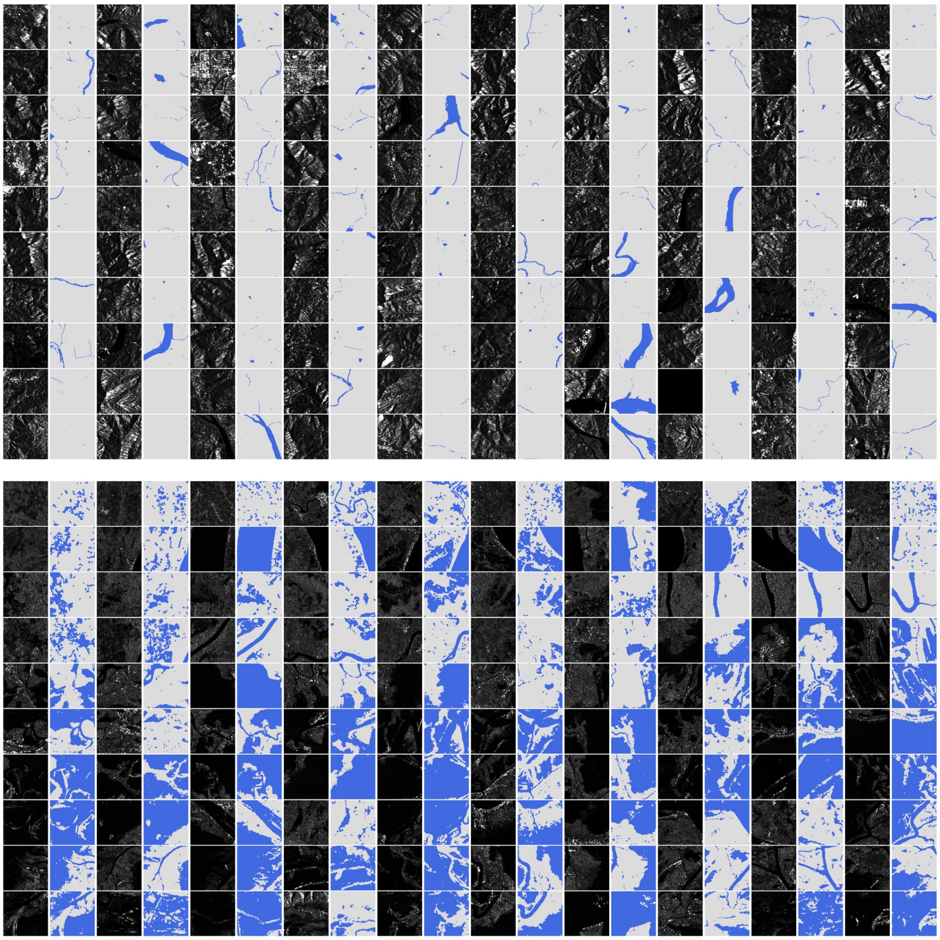

Figure 2.3 The examples of the final dataset (a) Lnadcover map based dataset (b) UNOSAT flood map based dataset.

3 Boundary-driven Adversarial Learning of Deep Neural Networks

The boundary-driven adversarial learning water segmentation network consists of a generator, a boundary distance map construction module, and a discriminator. The generator is a U-Net-based network that receives SAR images and segments water and non-water pixels. The boundary distance map construction module extracts the waterbody boundaries, and calculates the nearest distance from the boundaries. The discriminator is a modified simple classification network.

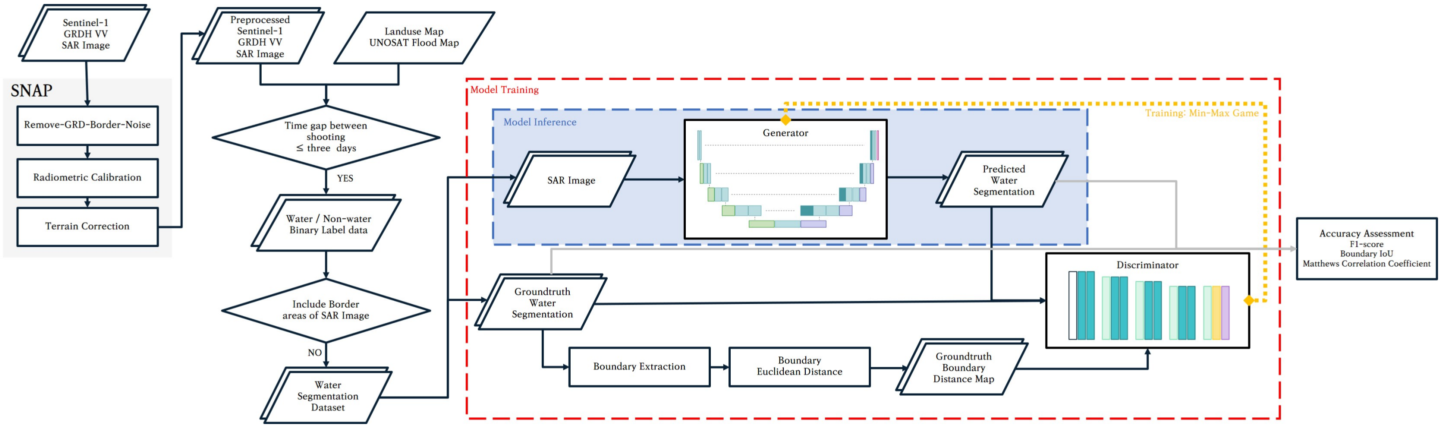

Figure 3.1 The full flowchart of the SAR water segmentation with adversarial learning.

3.1 Generator architecture

We designed our generative model of the SAR water segmentation network based on the U-Net architecture. The detailed structure of the generator is shown in Figure 3.2. The fundamental structure of U-Net is composed of the encoder and the decoder. Each encoder and decoder comprises four convolutional blocks, and in the middle of the encoder and the decoder, there is a bottleneck convolutional block with the deepest feature map. The convolutional block has two convolutional layers followed by Batch Normalization and ReLU activation with a strided convolutional layer. L1 kernel regularizer is added to convolutional layers to prevent overfitting. In contrast to the vanilla U-Net, we substituted max-pooling layers with strided convolutional layers with ReLU activation, and upsampling layers of the decoder with strided transpose convolutional layers with ReLU activation. With this replacements, non-trainable parameters of pooling layers change to trainable parameters, which improves and stabilizes the performance compared to the base model(Springenberg et al. 2014). Additionally, the strided convolutional layer has wider receptive fields even though the filter size is the same as pooling layers. This allows the model to increase its expressiveness and capture a larger context. The last convolution layer is for the segmentation, which is composed of the convolution layer with L1 regularizer, followed by sigmoid activation.

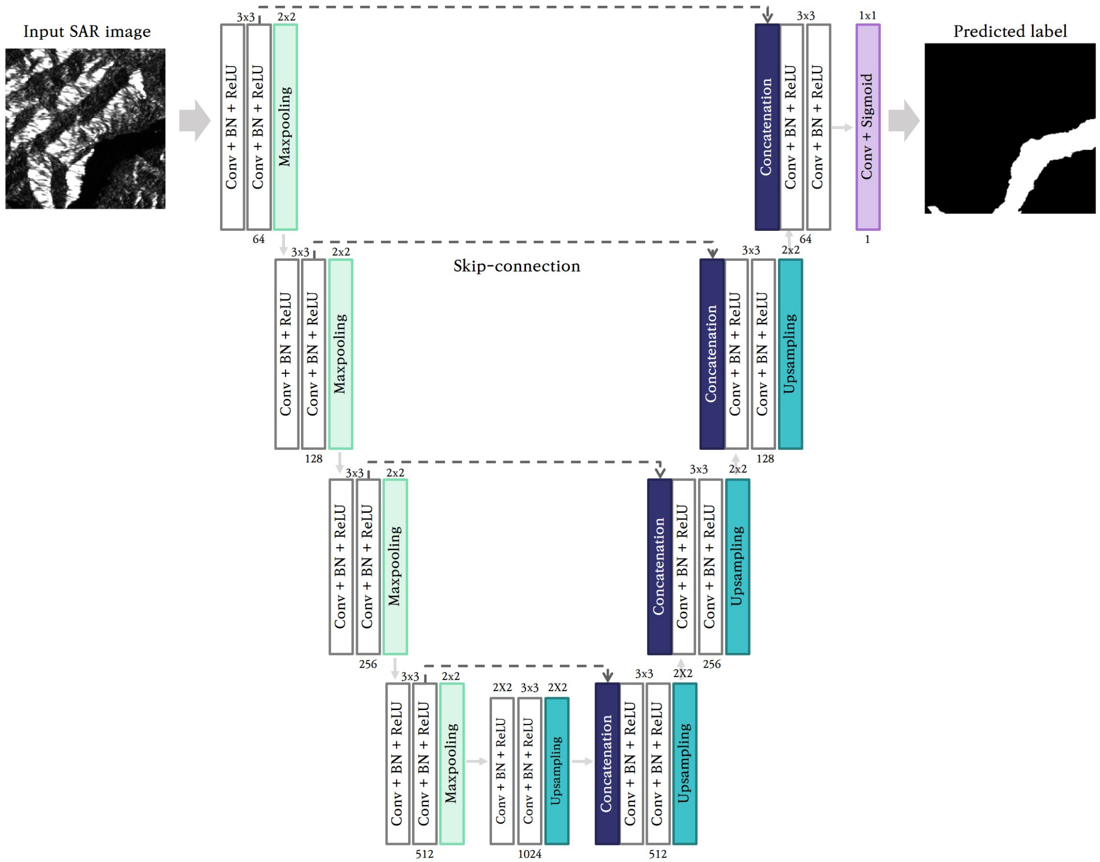

Figure 3.2 The detailed architecture of the SAR water segmentation generator. The number on the top of blocks means the kernel size of the 2D convolution window. The number on the bottom of blocks means the number of filters. The output of the second convolution layer for each encoder convolutional blocks is concatenate with the transpose convolution outputs.

3.2 Discriminator architecture

Figure 3.3 shows the discriminator architecture of the proposed model, guided by the boundary distance map of the groundtruth waterbody. The discriminator is composed of four convolutional blocks and one block for distinguishing the generated segmentation from the real groundtruth data. Each convolutional block consists of three convolutional layers. Every last convolutional layer is substituted with a strided convolutional layer to act as max-pooling layer but with trainable weights for stable performance(Springenberg et al. 2014). To prevent unstable training, we added a dropout layer for each convolutional block and employed Leaky ReLU as the activation function for every layer. The last block of the discriminator consists of a convolutional layer, global average pooling and a fully connected layer with a sigmoid activation function. This block discriminates the input image as either real groundtruth images or fake generated images. As the last activation function is sigmoid, the discriminator presents a probability score that indicates the likelihood that the input is real or generated.

The proposed discriminator is guided by one constraint: the boundary distance map based on the groundtruth label data. As the groundtruth label data is a binary image, high-performance edge detection algorithm like Canny edge detection are excessive and complex. Additionally, the extraction algorithm is activated every time the model learns. Therefore, a highly efficient boundary distance map is required.

We selected the boundary extraction algorithm based on morphology transformations, employing two morphological operations: erosion and dilation. Erosion shrinks from the boundaries of water bodies, whereas dilation expands to the boundaries of water bodies. The boundary is defined as the pixels that are removed or added with morphological operations. This morphological operation-based algorithm saves times by utilizing GPUs. Moreover, (Chen et al. 2002) verified that the morphological residue edge detector shows high performance with binary step edges. As the size of the structuring element determines the thickness of the boundary, and a thinner boundary is suitable for detailed performance, we set the size of the structuring element to three.

The boundary distance map represents the distance from each pixel to the nearest non-water pixel in a binary boundary image., calculated using the Euclidean distance. The procedure of calculating the boundary distance map is shown in Figure 3.4. As the values of distance fluctuate between images, we rescale it to a range of 0 to 1 to constrain it equally for every image in the discriminator. As shown in Figure 3.5, the border of the waterbody is emphasized in the boundary distance map, and the border of the shallow river is equally highlighted. As the segmentation task of the inner part is uncomplicated, this map could help the discriminator to concentrate on the boundary and small rivers for the distinguishing task. If the patch is fully composed of water or non-water, the boundary distance map is 1 for every pixel. Considering the variety of data that could be a constraint for the discriminator, we examined the contribution of the boundary distance map compared to the SAR image, the groundtruth boundary in Chapter 4.

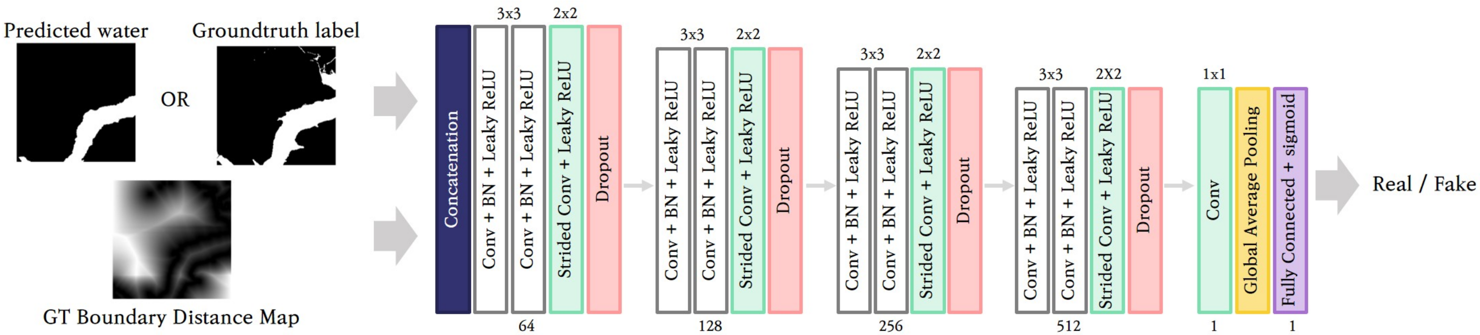

Figure 3.3 The detailed architecture of the SAR water segmentation discriminator. The number on the top of blocks means the kernel size of the 2D convolution window. The number on the bottom of blocks means the number of filters. The input of the discriminator is the output of SAR water segmentation generator or the groundtruth label data with boundary distance map extracted from the groundtruth label.

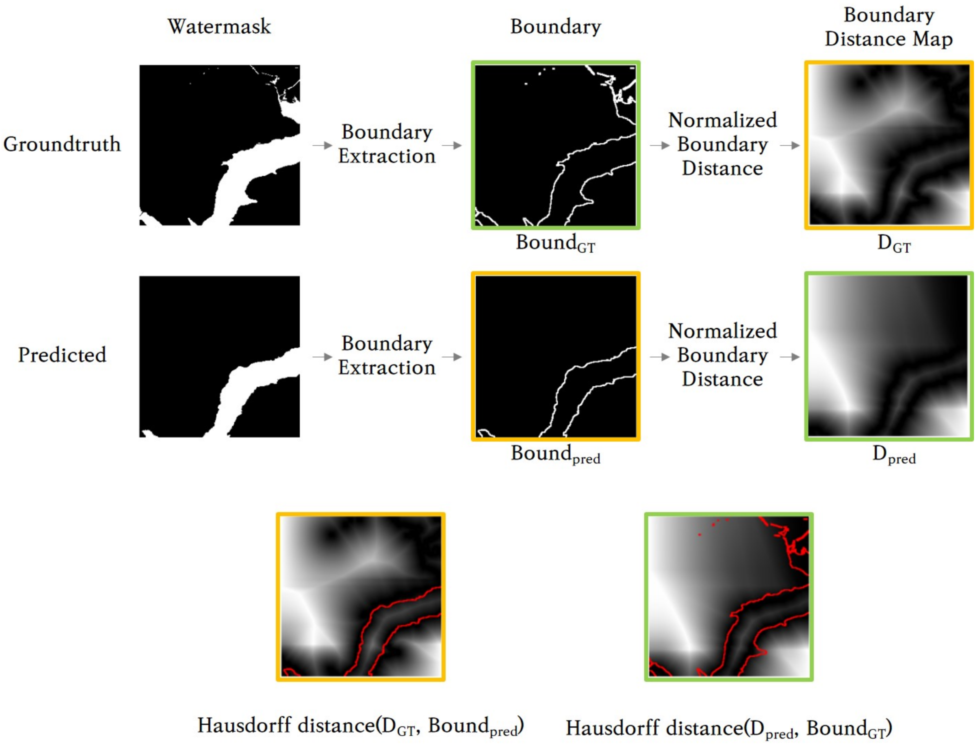

Figure 3.4 The procedure of constructing the boundary distance map.

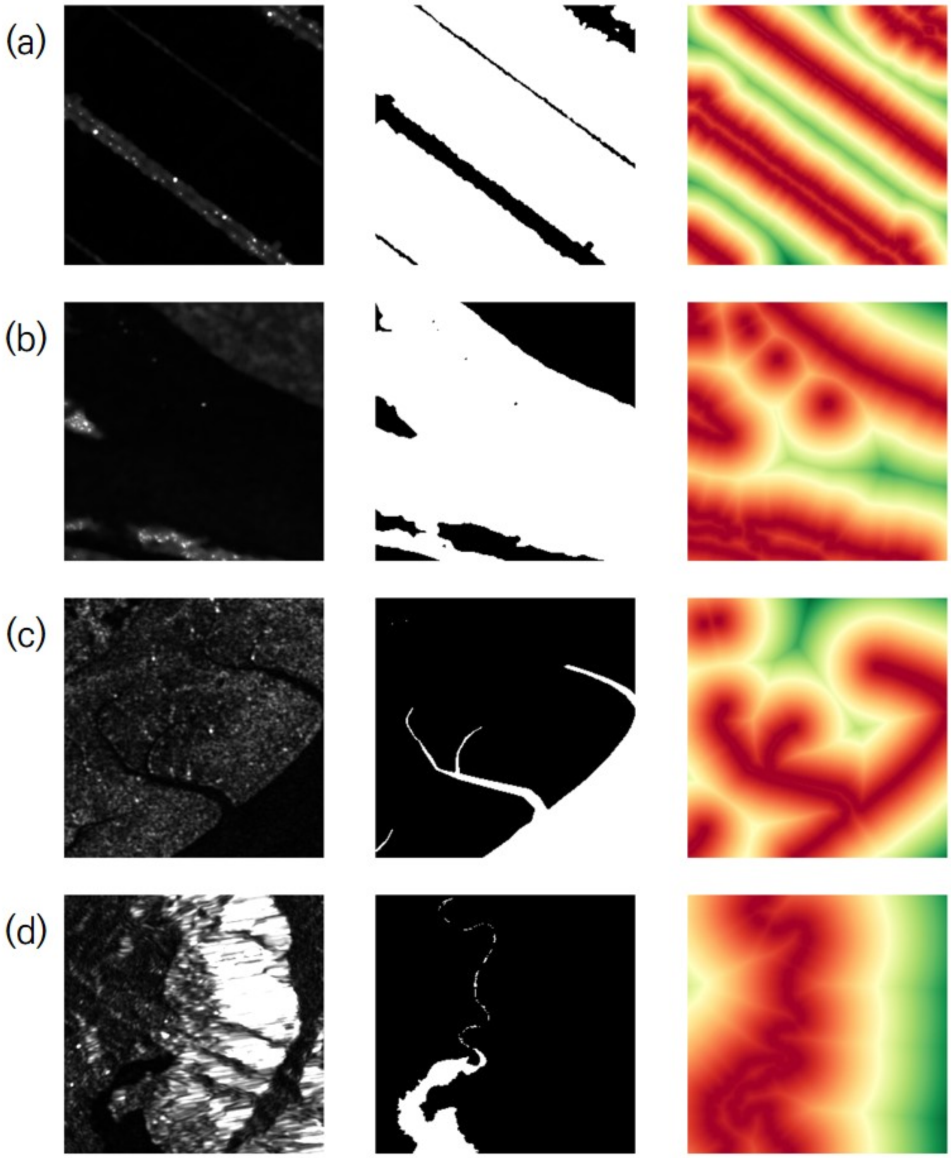

Figure 3.5 The examples of the boundary distance map with groundtruth label images.

3.3 Hybrid Loss

To attain high quality SAR water segmentation with precise boundaries, we propose our training loss as a weighted sum of three losses: Reconstruction loss, Bound loss, and adversarial loss.

$$L_{seg} = \lambda \cdot L_{Recon} + \alpha \cdot L_{Bound} + \beta \cdot L_{adv} \tag{3.1}$$

where, the λ, α, β are used for rescaling and balancing term between the reconstruction loss, bound loss and adversarial loss, respectively. The optimal value for the hyperparameters are compared in Table 4.4.

The reconstruction loss is a Binary Cross Entropy, which is widely used distribution-based loss for segmentation and pixel-level classification. Binary Cross Entropy(BCE) loss(Jadon no date) is the most commonly used distribution-based loss in binary classification and segmentation. It is defined as:

$$L_{BCE}(y, \hat{y}) = -\frac{1}{n}\sum_{i=1}^{n}[ylog(\hat{y}) + (1-y)log(1-\hat{y})] \tag{3.2}$$

where n is the number of the pixels of the predicted probability map, y is the ground truth label of each pixel, and $\hat{y}$ is the predicted probability of water.

The second loss function is the Hausdorff Distance loss(Karimi and Salcudean 2019). Hausdorff Distance(Crum et al. 2006) is defined as the maximum distance between a point in one of the two sets to its closest point in the other set. The Hausdorff Distance is a useful metric for evaluating the performance of segmentation, as it indicates the areas where the segmentation error is most significant. The bidirectional Hausdorff Distance between two pointsets X and Y is defined as:

$$HD(X, Y) = \max(\max_{y \in Y} \min_{x \in X} \parallel x - y \parallel_2, \quad \max_{x \in X} \min_{y \in Y} \parallel x - y \parallel_2) \tag{3.3}$$

However, since the Hausdorff Distance is fixed solely by the maximum error, directly application to the loss function leads to poor and unstable overall segmentation performance. Moreover, the sensitivity of the Hausdorff Distance to outliers and noise is well-known in computer vision. Nevertheless, (Karimi and Salcudean 2019) proposed the Hausdorff Distance loss as a method for training CNN-based segmentation based on distance transforms.

$$L_{Bound}(y, \hat{y}) = \frac{1}{|\Omega| + smooth} \sum_{\Omega} ((y - \hat{y})^2 \circ (d_y^\alpha + d_{\hat{y}}^\alpha)) \tag{3.4}$$

Here, Ω denotes the grid on which the image is defined, where the maximum of Ω is in accordance with all pixels. d_y and $d_{\hat{y}}$ denotes the normalized Euclidean distance transformation map of the predicted segmentation boundary and the ground truth segmentation boundary. respectively. As the value of the Euclidean distance transformation could vary across images, we rescaled the values to a range of 0 to 1. $\circ$ denotes the Hadamard product, and the hyperparameter α indicates how strongly we penalize large errors. As (Karimi and Salcudean 2019) verified, we fixed α to 1. Since the dataset may have images of full water or non-water, the value of Ω could be zero. To prevent a zero division error, we added the smooth term(1e-6) to the denominator.

The last loss function for the generator is the adversarial loss derived from the discriminator. The adversarial loss is defined by the Binary Cross Entropy loss. The goal of the generator is to generate images that are indistinguishable from the ground truth images. The adversarial loss optimizes the generator to fool the discriminator by setting up the discriminator with the generated image is correct(1) as follows:

$$L_{adv} = \sum L_{BCE}(D(G(x)|BDM(y)), 1) \tag{3.5}$$

where $D(X|Z)$ denotes the discriminator with the input X and the constraint Z, and $G(\bullet)$ denotes the generator. $BDM(y)$ denotes the boundary distance map based on the input y.

For the discriminator, the loss function is written as:

$$L_{disc} = \sum L_{BCE}(D(y|BDM(y)), 1) + \sum L_{BCE}(D(G(x)|BDM(y)), 0) \tag{3.6}$$

The goal of the discriminator is to distinguish between the real and predicted images. The discriminator loss measures how well the discriminator is able to differentiate the generated images(0) from the ground truth(1). As the adversarial loss has the value of 1 for the same input(the generated images), the generator and discriminator are trained simultaneously using a two-player minmax game.

4 Experiments

4.1 Experiment Settings

The training dataset of 28,569 sets is used without data augmentation. During training and testing, the input and output size of images is fixed into 256 * 256. Resizing by interpolation is not used for precise training and results. To prevent mixing of the dataset for training and testing, the dataset is segregated in a ratio of 80%:20%.

We utilized the Adam optimizer with an initial learning rate of 1e-6 for the generator and 1e-4 for the discriminator, with betas=(0.9, 0.999) for both. We added gradient clipping of 1.0 to converge faster than gradient descent with a fixed step size(Zhang et al. 2019). We set up the training for 30 epochs for ablation studies, but if there was no apparent disparity until 30 epochs, we continued training until 100 epochs. For comparison with other models, we trained the models for both 30 epochs and 100 epochs. If there were no improvements in validation generator loss for 10 epochs, the model saved the model with the minimum validation generator loss and finished the training. The batch size for each GPU is 16. The global batch size is determined by the multiplying the batchsize with the number of GPUs. The iterations for training and testing are 714, 178, respectively, when the number of GPUs is two and the global batch size is 32.

We implemented our model using Python 3.9. A ten-core PC with 40 Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz CPU and four GTX 3090 GPUs is used for both training and testing.

To verify the effectiveness of the proposed boundary-driven adversarial learning of deep neural networks, we used the same implementation settings and the dataset.

Table 4.1 Implementation details of experiments

Generator optimizer	Adam(1e-6, beta=0.9, beta2=0.999, clip_norm=1.0)
Discriminator optimizer	Adam(1e-4, beta=0.9, beta2=0.999, clip_norm=1.0)
Max Epoch	100 (30 for ablation studies)
Early Stopping	No improvements of validation generator loss for 10 epochs
Input data size	256 * 256
Global Batchsize	32
Train:test	8:2
Iterations	714: 178

4.2 Evaluation Metrics

We employ three measures to evaluate our method: F1-score(Sasaki 2007), Matthews correlation coefficient(Matthews 1975) and Boundary Intersection-over-Union(Boundary IOU)(Cheng et al. 2021).

F1-score is computed based on the confusion matrix, which contains true positive(TP), false positive(FP), true negative(TN), false negative(FN). F1-score is the harmonic mean of precision and recall and the result value varies in the interval [0, 1]. F1-score is calculated as follows:

$$precision = \frac{TP}{TP + FP} \tag{4.1}$$

$$recall = \frac{TP}{TP + FN} \tag{4.2}$$

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \tag{4.3}$$

F1-score is a popular and reliable metric, however, it could mislead if the dataset is balanced but with high FP or positively imbalanced(Chicco and Jurman 2020). In the field of the flood monitoring, false positive errors could misreport the inundation area and waste limited disaster response support. To complement this, we adopt another widely used metric, Matthews Correlation Coefficient(MCC). MCC is the discrete Pearson correlation coefficient based on the confusion matrix. MCC takes the value between [-1, 1]. 1 means the complete predictions, -1 means the complete opposite predictions and 0 represents the random predictions. Contrary to non-use of TN in F1-score, MCC weights all four elements of the confusion matrix equally. Therefore, MCC has high score only if the most positive and negative predictions are correctly predicted. MCC is written as:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{4.4}$$

The accuracy for region pixels and contour pixels does not always have strong positive correlation. The different growth rates when scaling up objects demonstrate that the precise metric for boundaries is additionally needed. Therefore, we adopt another boundary-based segmentation metric to precisely measure the segmentation quality with regard to boundaries. Boundary IoU(BIoU) (Cheng et al. 2021) denotes the Intersection-over-Union(IoU) between the boundary regions of the predicted segmentation image($\hat{y}$) and its ground truth mask(y). Given two images, BIoU extract the boundary mask that is within a distance d from each contour, and then computes the IoU of these two regions. BIoU is written as:

$$BIoU(Y, \hat{Y}) = \frac{|(Y_d \cap Y) \cap (\hat{Y_d} \cap \hat{Y})|}{|(Y_d \cap Y) \cup (\hat{Y_d} \cap \hat{Y})|} \tag{4.5}$$

where the boundary mask Y_d and $\hat{Y_d}$ are the pixel sets within d pixels from the groundtruth and predicted image contours respectively. BIoU overcome the problem of penalizing inner mask errors in Trimap IoU and also has a complementary relationship with mask-based F1-score. The parameter of distance from contour d controls the sensitivity of the measurement. We set d equals 2%, seven pixels distance for our dataset, as (Cheng et al. 2021) suggests.

4.3 Comparison to other segmentation models

For the quantitative evaluation, we compared our proposed model with two conventional deep learning segmentation models, namely U-Net and Deeplabv3+. All models were trained under the same implementation conditions, including the dataset, optimizer, and loss function. Table 4.2 shows the overall comparison results for both the maximum epochs of 30 and 100. Since the dataset for SAR water segmentation is substantial, all models also exhibit satisfactory performance. However, the quantitative results demonstrate that our proposed model consistently outperforms the other segmentation models in terms of every evaluation metric and at both epochs of training. Our proposed model achieves superior performance at the maximum epoch of 30 compared to the other models at the maximum of epoch of 100, indicating the efficiency of our proposed model during training. The gap between F1-score and MCC of our proposed model is 3.86% and 3.43% at the maximum epoch of 30 and 100, respectively, which is smaller than both U-Net at 5.44% and 4.26%, and Deeplabv3+ at 4.96% and 4.33%. As MCC also considers negative predictions, the proposed model also focuses on reducing the false errors than other models. Furthermore, the disparity in BIoU at 4.71% and 3.32%, which is larger than in other evaluation metrics, with MCC at 3.24% and 1.73% and the F1-score at 1.65% and 0.92%. This demonstrates that Boundary-aware Adversarial Learning is indeed competent in detecting accurate boundaries.

Figure 4.2 displays the visualization comparison of the three models. (a) and (b) represents the area of narrow rivers. The proposed model more elaborately segments the narrow rivers and boundaries than the other models. (c) and (d) depicts mountainous regions, where the false positives easily occur by radar shadow. In (d), all three model predict correctly into non-water; however, the U-Net model misjudges the radar shadow as water. In (e), the proposed model yields the closest prediction to the ground truth for a small reservoir, while the other models underestimate its area. Additionally, the proposed model demonstrates effectiveness in detecting small and fine regions, as evident in (f) to (h), where small objects are present amidst other classes.

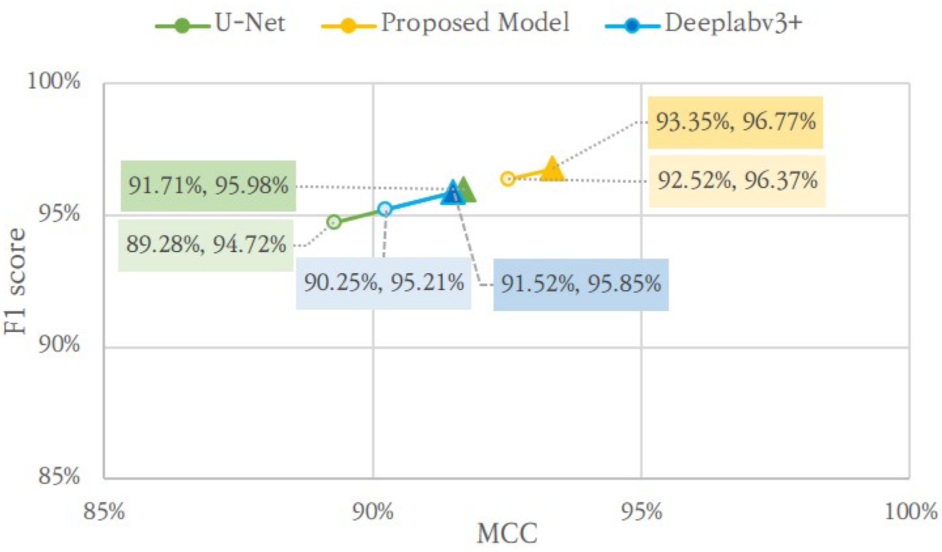

Figure 4.1 Illustration of F1-score and MCC on three models: U-Net, Deeplabv3+ and the proposed model. The brighter color represents the result of max epoch 30, and vivid color represents the result of max epoch 100.

Table 4.2 Performance comparison of SAR Water Segmentation. Best results are shown in bold

Model	Max Epoch 30			Max Epoch 100		
	BIoU	MCC	F1-score	BIoU	MCC	F1-score
U-Net	74.48	89.28	94.72	78.17	91.71	95.98
DeepLabV3+	74.81	90.25	95.21	76.89	91.52	95.85
Proposed Model	**79.19**	**92.52**	**96.37**	**80.21**	**93.35**	**96.77**

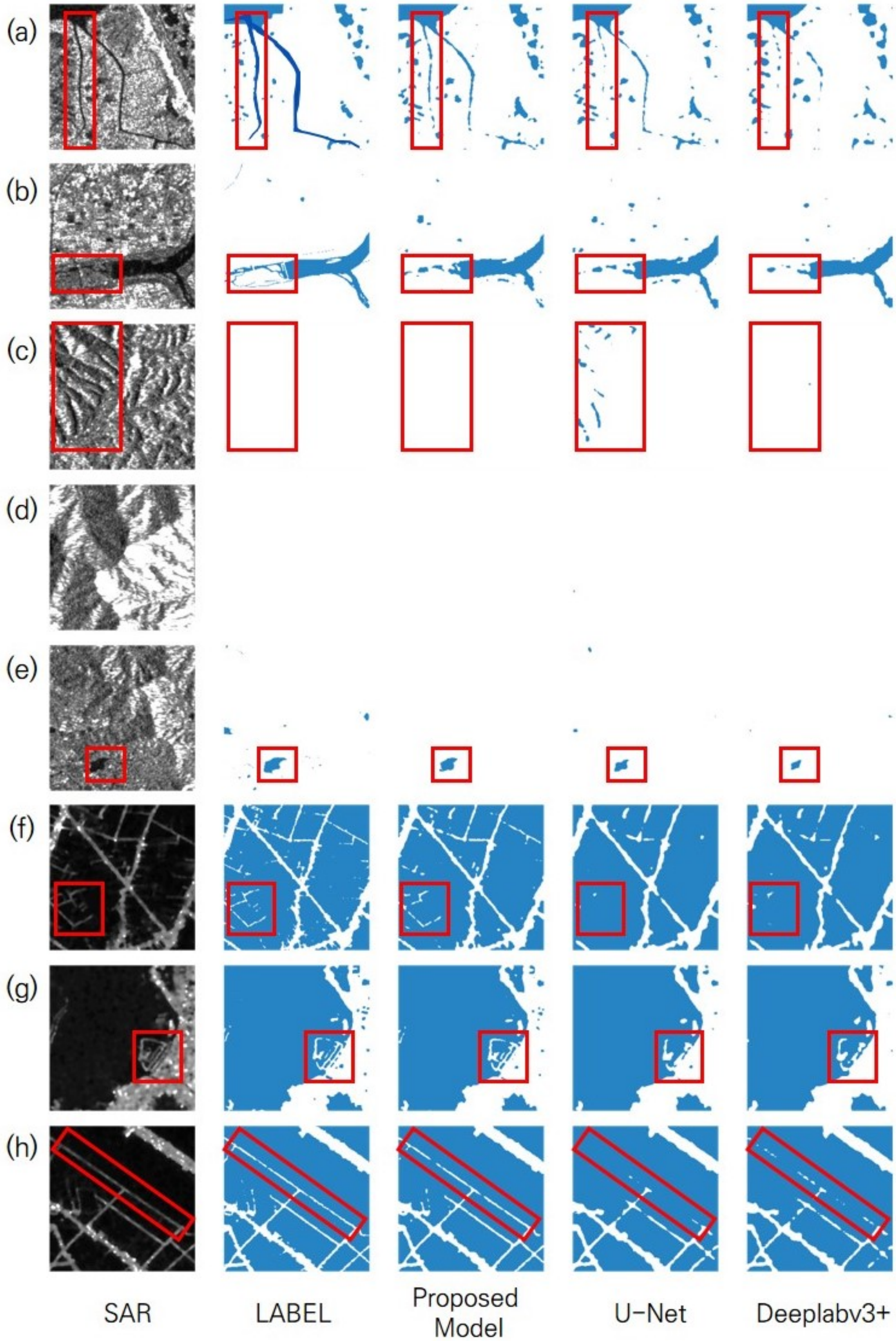

Figure 4.2 Qualitive performance evaluation of the SAR Water Segmentation. From left to right, the SAR image, the ground truth label, the output of the proposed model, vanilla U-Net, DeepLabv3+ is shown.

4.4 Ablation Studies

We conduct a series of ablation studies to analyze the impact of alterations on model and loss functions of the proposed boundary-driven adversarial learning of deep neural networks. There are three ablation studies: ablation on adversarial learning and constraints, ablation on hybrid loss and weight parameters, and ablation on constitution with strided convolutional layer.

Ablation on Adversarial Learning and Constraints

To prove the effectiveness of adversarial learning on the SAR water segmentation task, we compared the stand-alone U-Net model (without discriminator) and with the one using the discriminator along with other constraints. Table 4.3 reports that incorporating the discriminator into the SAR water segmentation increases BIoU by 3.9%, MCC by 3.76% and F1-score by 1.81% on average. We replaced our ground truth boundary distance map with SAR images and ground truth boundary, which resulted in slight improvements of MCC and F1-score with a decrease in BIoU. Direct input of the ground truth boundary shows the highest BIoU but with lower MCC and F1-score. It indicates that explicit use of the border information biased the model towards boundaries and leads to the model failing to see the overall context.

Table 4.3 Ablation study on discriminator and constraints

Discriminator	Constraint	BIoU	MCC	F1-Score
without discriminator		74.28	88.22	94.29
with discriminator	SAR	78.21	91.97	96.10
	GT Boundary	78.46	91.92	96.08
	GT Boundary Distance Map	77.87	92.05	96.12

Ablation on Loss Function and Rescaling Parameters

Loss function was examined from three points of view: the optimum reconstruction loss function, rescaling parameters of bound loss and the addition of Hausdorff Distance loss. The candidates for the reconstruction loss are Binary Cross Entropy loss, Focal Cross Entropy loss and Dice loss. To match the scale, λ is fixed to 0.0001 for Cross Entropy losses and 1 for Dice loss, which converges the range into [0,1]. Table4.4 demonstrates that the Binary Cross Entropy loss outperforms other loss functions by at least 6.04% in terms of BIoU, 3.25% in terms of MCC and 1.92% in terms of F1-score.

To validate the effectiveness of Hausdorff Distance loss, the optimum α needed to be analyzed in advance. The rescaling parameters of the losses are λ, α and β for the reconstruction loss, bound loss and adversarial loss. To rescale, we scrutinized the ranges for each loss function and determined the values. In the Figure 4.3 (a), as the value of the Binary Cross Entropy loss mostly ranges from [1e+3, 1e+4] for our model, we set λ and β to 0.0001 to rescale it to [1e-1, 1]. The value of the Hausdorff Distance loss for our model mostly ranges in [0, 10] in Figure 4.3 (b), so the default value of α is set to 0.01 to rescale it to [0.01, 0.1]. As the Hausdorff Distance loss has fluctuating values and outliers, the range of the bound loss is smaller than that of the reconstruction loss for stable training. The beta is fixed at 0.0001 as it has same loss function(Binary Cross Entropy) as the Reconstruction loss. To fine-tune the rescaling parameters, we experimented with different values of alpha, such as 0.1 and 0.001. As a result, $\alpha = 0.001$ is determined to be the optimal value because the BIoU and MCC have lower values, but the improvement of F1-score is much larger.

Finally, the addition of the Hausdorff Distance loss with optimal rescaling is compared without bound loss. We found that MCC is 0.14% lower in using the bound loss, but the improvements of BIoU and F1-score with 0.09% and 1.48% are judged to be more meaningful values.

These ablation studies indicate that when the stable reconstruction loss is supported, the bound loss can accomplish the task with complete attention. If the fluctuating bound loss takes on a significant position($\alpha = 0.1$), the accuracy of the boundary is rather declined. Even though the Dice loss is also a popular metric and has strength on imbalanced datasets, the gradient for backpropagation is more unstable than that of the Binary Cross Entropy in our task. As training GANs is intrinsically unstable, even little alternating variation of the reconstruction loss may have a critical influence. The Focal Cross Entropy is originally designed for handling class imbalance, but as our data has a variety of water ratios, the Focal Cross Entropy shows the worst evaluation metrics.

Table 4.4 Ablation study on different reconstruction loss functions and the loss rescaling parameters

L_{Recon}		L_{Bound}		L_{adv}		BIoU	MCC	F1-score
Loss	λ		α		β			
BCE	0.0001	-		BCE	0.0001	77.87	92.05	96.12
	0.0001		0.001			77.96	91.91	97.60
BCE	0.0001		0.1			77.58	91.82	96.00
	0.0001	HD	0.01	BCE	0.0001	78.03	91.92	96.06
FCE	0.0001		0.01			67.11	86.28	92.53
Dice	1		0.01			71.99	88.67	94.14

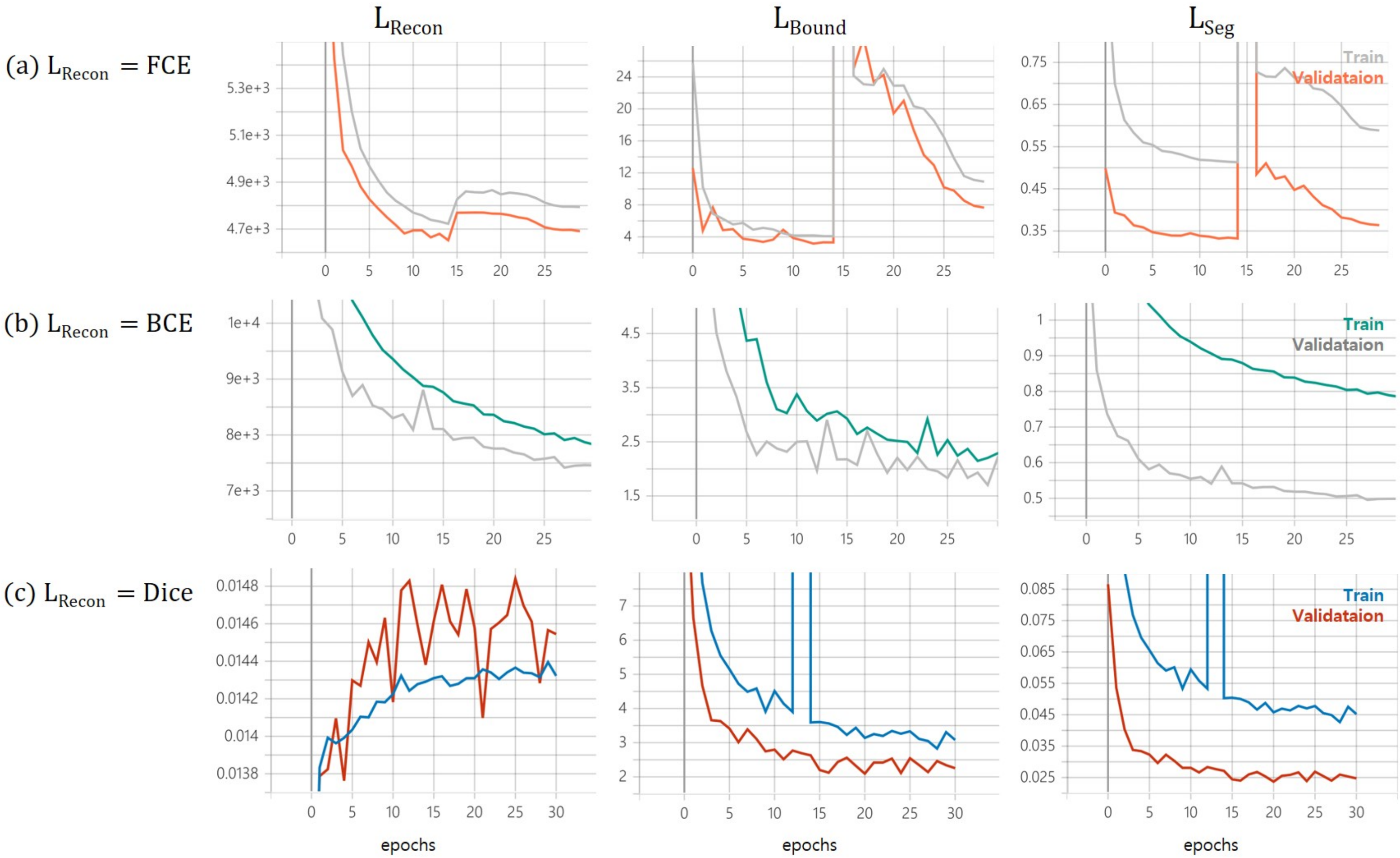

Figure 4.3 Illustration of different reconstruction loss functions(L_{Recon}) and corresponding bound loss(L_{Bound}) and total segmentation loss(L_{Seg}) for each training epoch.

Ablation on strided Convolutional Layers

The effect of substituting pooling layers with strided convolutional layers is shown in Table 4.5. We compared the architectures under two conditions: Max epoch of 30 and 100, while changing the bound loss scaling parameter α. When α is set to 0.01 and the max epoch is 30, U-Net with strided convolutional layers outperforms the original U-Net. However, when it comes to the max epoch of 100, U-Net with strided convolutional layers performs lower than the original U-Net. This indicates that the replacement with strided convolutional layers seems to accelerate training, but as training progresses, the model suffers from overfitting. As the replacement of pooling layers increases the complexity of the model, the model fails to generalize. Therefore, we also experimented with a different α value of 0.001. In this case, except for the F1-score in Max epoch 30, the vanilla U-Net demonstrates the competence than the model with strided convolutional layers.

Table 4.5 Ablation study of the Model Architecture on different bound loss rescaling parameter α

α	Architecture	Max epoch 30			Max epoch 100		
		BIoU	MCC	F1-score	BIoU	MCC	F1-score
0.01	U-Net	74.86	89.41	94.83	79.89	93.25	96.70
	U-Net with strided conv	78.26	92.04	96.13	78.76	92.42	96.32
0.001	U-Net	79.19	92.52	96.37	80.21	93.35	96.77
	U-Net with strided conv	77.96	91.91	97.60	79.02	92.55	96.37

5 Discussion

To assess its applicability in practical scenarios, we conducted experiments to evaluate the performance of the proposed model with entirely new images. Since the model is trained on patches with 256 * 256 pixels, it is not directly applicable to large scene SAR images in practical monitoring, which typically have dimensions of approximately 20000 * 30000 pixels. To prevent memory errors, we divide the large scene SAR images into patches of the same size as the training dataset. To maintain continuity and precision in the predicted images, we split the large scene SAR images into overlapping patches and then reassemble them into large scene SAR images by averaging the probabilities obtained from the individual patches. This approach ensures that the proposed model can be effectively applied to large-scale SAR images used in real-world flood monitoring applications.

Using the overlapping prediction, we segmented water in six large scene SAR images of Korea Peninsula. The detailed list of SAR images is shown in Table 5.1. Since there are no ground truth label for these images, we compare the proposed model with that of the SAR images, Vanilla U-net and DeepLabv3+. The visualization comparisons of the three models are displayed in Figure 5.1. It can be observed that the proposed model outperforms than other models. First, False Positive(FP) errors easily arise in areas where the surface is flat enough to exhibit specular reflection properties, such as golf courses. In (a) and (g), U-Net shows plenty of False Positive errors in the golf course areas. In contrast, DeepLabv3+ successfully avoids detecting the golf courses as water, but it fails to detect the real river. On the other hand, the proposed model shows few false positives and accurately predicts the presence of the real river. Second, false positives also occur in area where the radar signal does not reach the ground due to obstructions, resulting in radar shadow appearing as dark regions in the SAR images. In the mountainous regions, (b) and (h), U-Net significantly detects the radar shadow as water. Similarly, DeepLabv3+ shows few false positives but considerable false negatives. The proposed model achieves a better balance with some false positives but highly accurate true positives. Lastly, we also inspected that our proposed model consistently outperforms other models regardless of the river width. In (c) and (d), large rivers with widths of 0.5-1km are shown, and as large rivers usually exhibit simple boundaries with apparent signals, all models correctly detect them. In (j), a river with a width of 150m is displayed, and the waterbody is well detected except for the border area. For different cases of meandering rivers with widths of 60-70m in (d), (e), and (k), DeepLabv3+ performs the least accurately when it comes to smaller rivers. On the other hand, U-Net shows better results for smaller rivers. In the case of small reservoirs like (f) and (j), the other two models underperform in capturing the details of the reservoirs or fail to detect the small-scale reservoirs. These results indicate the potential of the proposed model for practical application in real-world flood monitoring scenarios, especially considering the accurate segmentation of narrow and intricate waterbodies and little false positives.

The continuity of precision regardless of the incidence angle is also verified. Figure 5.2 displays the results at two different incidence angles. We can obtain SAR images with different incidence angles at the same location due to the overlapping swaths of adjacent orbits of Sentinel-1. AS two SAR images has a 7-day interval with same orbit direction, the overlapping area is the optimal test site to assess the effects of the incidence angle. As observed in Figure 5.2 (b) and (c), the results of the vanilla U-Net and the DeepLabv3+ are influenced by the incidence angles, particularly in the radar shadow zone. At a mean incidence angle of 36°, which is relatively small (Figure 5.2 (b)), all models accurately predict water. In contrast, when the incidence angle increases to 45° (Figure 5.2 (c)), false positive errors are more likely to occur due to the decrease in signal strength. The vanilla U-Net and DeepLabv3+ have many false positives in the radar shadow zone. However, the proposed model is relatively immune from incidence angle differences. This experiment also indicates that the proposed model is suitable for practical flood monitoring applications.

However, the model still has limitations in detecting water in urban areas. In the middle of the city, plenty of high reflectance signals are intertwined with the area of the water, resulting in higher reflectance values for urban river pixels compared to typical waterbody pixels. Therefore, all of the models have difficulty in detecting urban rivers, including large rivers. This challenge can potentially be overcome by applying geospatial layers with SAR images, but auxiliary data construction is additionally required for training and practical implementation. As urban rivers are visually apparent, future research using contrast and shape information could potentially overcome this problem.

Table 5.1 The detailed information of Sentinel-1 SAR images for application test in Korean Peninsula

Satellites	Beam Mode, Processing Level	Acquisition Date(UTC)	Orbit Direction
Sentinel-1A	IW, GRDH	2022-08-16T09:24:22 - 09:24:47	ASCENDING
Sentinel-1A	IW, GRDH	2022-08-16T09:23:57 - 09:24:22	ASCENDING
Sentinel-1A	IW, GRDH	2022-08-16T09:23:28 - 09:23:57	ASCENDING
Sentinel-1A	IW, GRDH	2022-08-09T09:32:25 - 09:32:50	ASCENDING
Sentinel-1A	IW, GRDH	2022-08-09T09:32:00 - 09:32:25	ASCENDING
Sentinel-1A	IW, GRDH	2022-08-09T09:31:30 - 09:32:00	ASCENDING

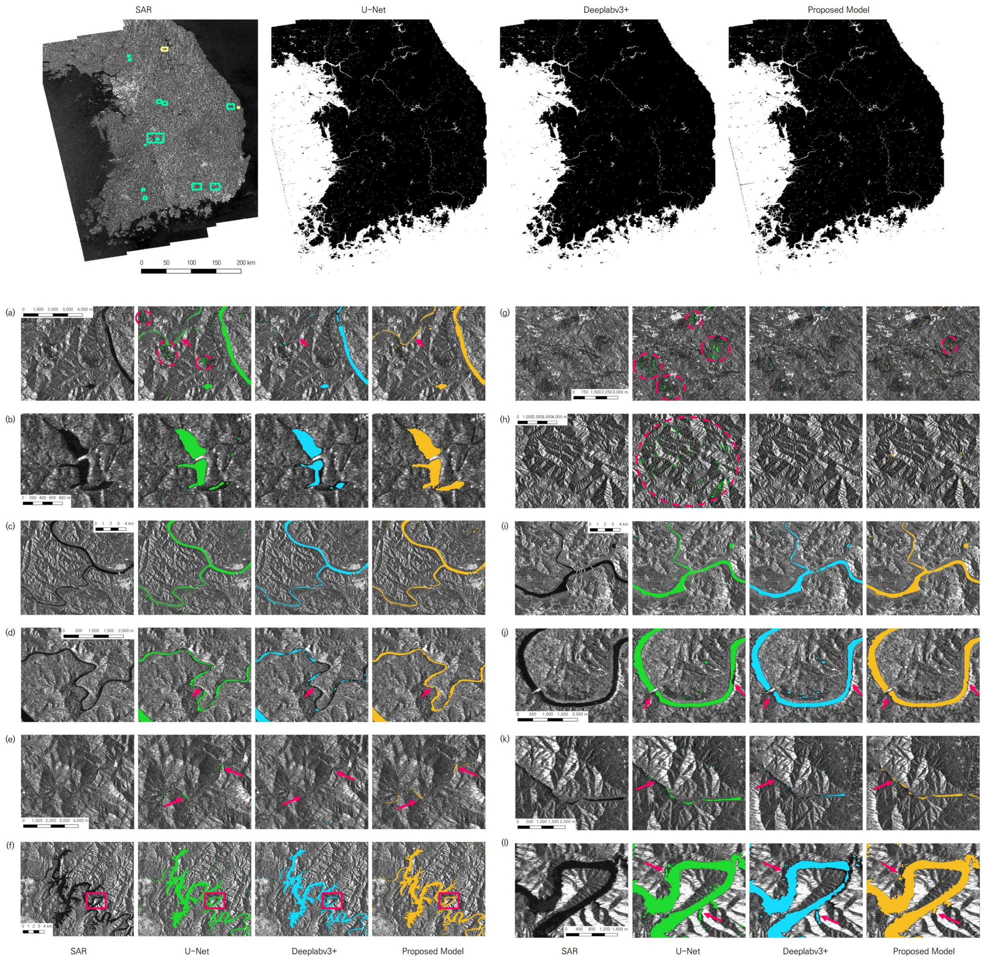

Figure 5.1 Application results in Korea Peninsula of the waterbodies monitoring. The green, blue, yellow color represents the result of U-Net, Deeplabv3+, and the proposed model, respectively.

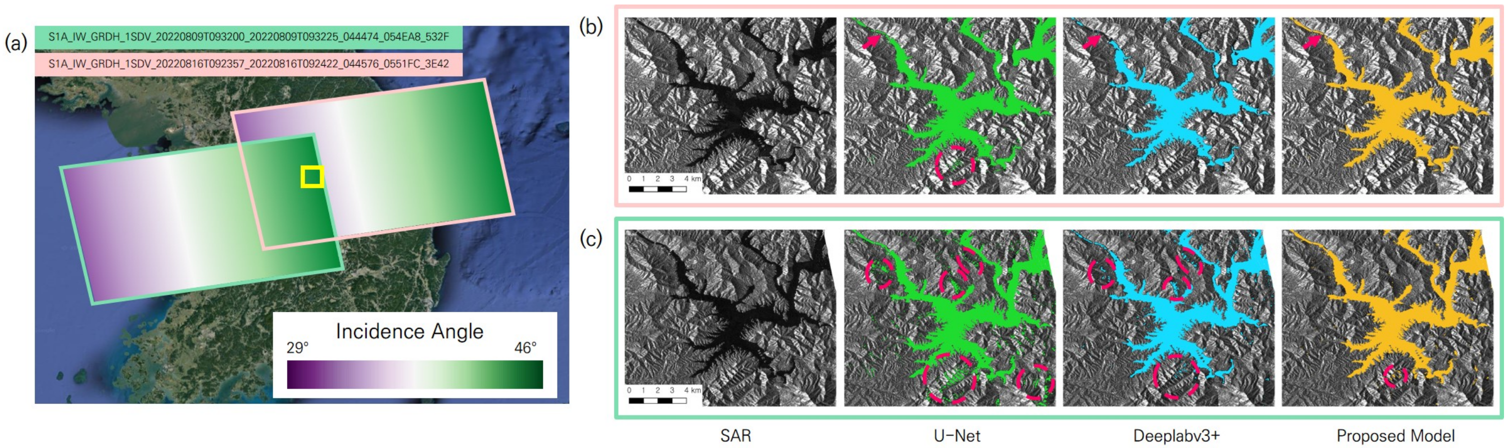

Figure 5.2 The results of the three models by the incidence angle difference. (a) shows the incidence angle of two large scene SAR images: `S1A_IW_GRDH_1SDV_20220809T093200_20220809T093225_044474_054EA8_532F`(mint) and `S1A_IW_GRDH_1SDV_20220816T092357_20220816T092422_044576_0551FC_3E42`(pink). The yellow box is the region of interest where two images overlap with different incidence angles. (b) and (c) show the area of the yellow box in the pink-colored SAR image and mint-colored SAR image, respectively. The green, blue, yellow color represents the result of U-Net, Deeplabv3+, and the proposed model, respectively.

6 Conclusion

Inundation monitoring is crucial for mitigating the impacts of flooding on human life, infrastructure, and the environment. For accurate flood monitoring, preciseness in water segmentation needs to be supported. As SAR is all-weather and available day and night, water segmentation using SAR is actively progressing. Although the application of CNN in SAR water segmentation has advanced the accuracy and efficiency, there are still deficiencies when utilizing it in practice. Waterbodies lookalike areas, such as golf courses or dark regions due to radar shadows, cause challenges. Moreover, since the number of pixels for narrow rivers and boundaries is fewer than the number of waterbodies' pixels, the vanilla CNN models inevitably penalize the small rivers and river boundaries. These difficulties may be particularly exacerbated when training data is scarce.

To address this, we construct the sufficient training dataset using UNOSAT flood map and the landuse map of Korea. In addition, we firstly propose the use of adversarial training from GAN in SAR water segmentation to improve the accuracy even in fine details. Adversarial training involves two neural networks, a generator and a discriminator. The generator generates fake samples that are indistinguishable from real samples, while the discriminator attempts to correctly distinguish between real and fake samples. Through learning from each other, both models improve over time, which is why this process is called adversarial training.

We construct the Boundary-aware SAR water segmentation model with adversarial training, and modify the architecture and loss functions to emphasize boundaries and narrow rivers. Specifically, adversarial training with the boundary distance map enforces boundary detection and reduces False Positives. The Hausdorff Distance loss helps our model to detect detailed waterbodies regardless of scale.

As a result, our proposed model outperforms other segmentation models in all evaluation metrics such as Boundary IoU, MCC, F1-score. Ablation studies demonstrate the potential of adversarial learning and the constraint for the discriminator, and also verify the optimality of the parameters and model architecture. The hybrid loss of Binary Cross Entropy loss, Hausdorff Distance loss and adversarial loss with optimal rescaling weights keeps training stable and efficient. The experiment with completely new SAR images of Korean Peninsula also proves the possibility of applying our model to practical flood monitoring. This experiment demonstrates that the model properly detects the water regardless of differences in the size of images and acquisition conditions.

We demonstrates that the Boundary-aware SAR water segmentation model with adversarial training learned the optimal formula for precisely and practically detecting water, including borders and narrow rivers, but without false positives. Our proposed model can be efficiently applied to near-real-time flood monitoring, where the construction of auxiliary data is not necessary. Furthermore, the precisely detected water boundaries can serve as fundamental data for water level estimation using SAR.

Bibliography

Chen, Liang-Chieh, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam, 2018: Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European conference on computer vision (ECCV)*, 801–818.

Chen, Ting, QH Wu, Reza Rahmani-Torkaman, and Jim Hughes, 2002: A pseudo top-hat mathematical morphological approach to edge detection in dark regions. *Pattern Recognition*, **35**, 199–210. ISSN: 0031-3203.

Cheng, Bowen, Ross Girshick, Piotr Dollár, Alexander C Berg, and Alexander Kirillov, 2021: Boundary iou: improving object-centric image segmentation evaluation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 15334–15342.

Chicco, Davide, and Giuseppe Jurman, 2020: The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. *BMC genomics*, **21**, 1–13.

Crum, William R, Oscar Camara, and Derek LG Hill, 2006: Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE transactions on medical imaging*, **25**, 1451–1461. ISSN: 0278-0062.

Denbina, Michael, Zaid J Towfic, Matthew Thill, Brian Bue, Neda Kasraee, Annemarie Peacock, and Yunling Lou, 2020: Flood mapping using uavsar and convolutional neural networks. *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 3247–3250. ISBN: 1728163749.

Dong, Shan, Long Pang, Yin Zhuang, Wenchao Liu, Zhanxin Yang, and Teng Long, 2019: Optical remote sensing water-land segmentation representation based on proposed sns-cnn network. *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 3895–3898. ISBN: 153869154X.

Fan, Jianchao, and Chuan Liu, 2023: Multitask gans for oil spill classification and semantic segmentation based on sar images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **16**, 2532–2546. DOI: 10.1109/JSTARS.2023.3249680.

Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, 2020: Generative adversarial networks. *Communications of the ACM*, **63**, 139–144. ISSN: 0001-0782.

Guo, Zhishun, Lin Wu, Yabo Huang, Zhengwei Guo, Jianhui Zhao, and Ning Li, 2022: Water-body segmentation for sar images: past, current, and future. *Remote Sensing*, **14**, 1752. ISSN: 2072-4292.

Höfle, Bernhard, Michael Vetter, Norbert Pfeifer, Gottfried Mandlburger, and Johann Stötter, 2009: Water surface mapping from airborne laser scanning using signal intensity and elevation data. *Earth Surface Processes and Landforms*, **34**, 1635–1649. ISSN: 0197-9337.

Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros, 2017: Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125–1134.

Jadon, Shruti, no date: A survey of loss functions for semantic segmentation. *2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB)*. IEEE, 1–7. ISBN: 1728194687.

Kang, Wenchao, Yuming Xiang, Feng Wang, Ling Wan, and Hongjian You, 2018: Flood detection in gaofen-3 sar images via fully convolutional networks. *Sensors*, **18**, 2915. ISSN: 1424-8220.

Karimi, Davood, and Septimiu E Salcudean, 2019: Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Transactions on medical imaging*, **39**, 499–513. ISSN: 0278-0062.

Kim, Junwoo, Hwisong Kim, Hyungyun Jeon, Seung-Hwan Jeong, Juyoung Song, Suresh Krishnan Palanisamy Vadivel, and Duk-jin Kim, 2021: Synergistic use of geospatial data for water body extraction from sentinel-1 images for operational flood monitoring across southeast asia using deep neural networks. *Remote Sensing*, **13**, 4759. ISSN: 2072-4292.

Klemenjak, Sascha, Björn Waske, Silvia Valero, and Jocelyn Chanussot, 2012: Automatic detection of rivers in high-resolution sar data. *IEEE Journal of selected topics in applied earth observations and remote sensing*, **5**, 1364–1372. ISSN: 1939-1404.

Lei, Baiying, Zaimin Xia, Feng Jiang, Xudong Jiang, Zongyuan Ge, Yanwu Xu, Jing Qin, Siping Chen, Tianfu Wang, and Shuqiang Wang, 2020: Skin lesion segmentation via generative adversarial networks with dual discriminators. *Medical Image Analysis*, **64**, 101716. ISSN: 1361-8415.

Li, Donghui, Jia Liu, Fang Liu, Wenhua Zhang, Andi Zhang, Wenfei Gao, and Jiao Shi, 2022: A dual-fusion semantic segmentation framework with gan for sar images. *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 991–994. DOI: `10.1109/IGARSS46834.2022.9884931`.

Li, Yongqing, Xinrong Lyu, Alejandro C. Frery, and Peng Ren, 2021: Oil spill detection with multiscale conditional adversarial networks with small-data training. *Remote Sensing*, **13**. ISSN: 2072-4292. DOI: `10.3390/rs13122378`. URL: `https://www.mdpi.com/2072-4292/13/12/2378`.

Liu, H, and KC Jezek, 2004: Automated extraction of coastline from satellite imagery by integrating canny edge detection and locally adaptive thresholding methods. *International journal of remote sensing*, **25**, 937–958. ISSN: 0143-1161.

Liu, Zhongling, Fei Li, Ning Li, Robert Wang, and Heng Zhang, 2016: A novel region-merging approach for coastline extraction from sentinel-1a iw mode sar imagery. *IEEE Geoscience and remote sensing letters*, **13**, 324–328. ISSN: 1545-598X.

Long, Jonathan, Evan Shelhamer, and Trevor Darrell, 2015: Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431–3440.

Matthews, Brian W, 1975: Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, **405**, 442–451. ISSN: 0005-2795. URL: `https://www.sciencedirect.com/science/article/pii/0005279575901099`.

Mirza, Mehdi, and Simon Osindero, 2014: Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Pai, Manohara MM, Vaibhav Mehrotra, Shreyas Aiyar, Ujjwal Verma, and Radhika M Pai, 2019: Automatic segmentation of river and land in sar images: a deep learning approach. *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*. IEEE, 15–20. ISBN: 1728114888.

Park, Hyenok, Hong Joo Lee, Hak Gu Kim, Yong Man Ro, Dongkuk Shin, Sa Ra Lee, Sung Hoon Kim, and Mikyung Kong, 2019: Endometrium segmentation on transvaginal ultrasound image using key-point discriminator. *Medical physics*, **46**, 3974–3984. ISSN: 0094-2405.

Pradhan, Biswajeet, Maher Ibrahim Sameen, and Bahareh Kalantar, 2017: Optimized rule-based flood mapping technique using multitemporal radarsat-2 images in the tropical region. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **10**, 3190–3199. ISSN: 1939-1404.

Ronci, Federico, Corrado Avolio, Mauro Di Donna, Massimo Zavagli, Veronica Piccialli, and Mario Costantini, 2020: An adversarial learning approach for oil spill detection from sar images. *2020 IEEE Radar Conference (RadarConf20)*, 1–4. DOI: `10.1109/RadarConf2043947.2020.9266475`.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox, 2015: U-net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 234–241. ISBN: 3319245732.

Sasaki, Yutaka, 2007: The truth of the f-measure. *Teach Tutor Mater*.

Silveira, Margarida, and Sandra Heleno, 2008: Water/land segmentation in sar images using level sets. *2008 15th IEEE International Conference on Image Processing*. IEEE, 1896–1899. ISBN: 1424417651.

Springenberg, Jost Tobias, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller, 2014: Striving for simplicity: the all convolutional net. *arXiv preprint arXiv:1412.6806*.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin, 2017: Attention is all you need. *Advances in neural information processing systems*, **30**.

White, Lori, Brian Brisco, Mohammed Dabboor, Andreas Schmitt, and Andrew Pratt, 2015: A collection of sar methodologies for monitoring wetlands. *Remote sensing*, **7**, 7615–7645. ISSN: 2072-4292.

Xie, Lei, Hong Zhang, and Chao Wang, 2015: Water-body types classification using radarsat-2 fully polarimetric sar data. *2015 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES).* IEEE, 1–5. ISBN: 1467377147.

Yuan, Kunhao, Xu Zhuang, Gerald Schaefer, Jianxin Feng, Lin Guan, and Hui Fang, 2021: Deep-learning-based multispectral satellite image segmentation for water body detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **14**, 7422–7434. ISSN: 1939-1404.

Zhang, Jingzhao, Tianxing He, Suvrit Sra, and Ali Jadbabaie, 2019: Why gradient clipping accelerates training: a theoretical justification for adaptivity. *arXiv preprint arXiv:1905.11881.*

Zhou, Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang, 2018: Unet++: a nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, 3–11.

Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A Efros, 2017: Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE international conference on computer vision*, 2223–2232.

초 록

기후변화가 가속화로 인해 수재해의 빈도와 강도 예측이 어려워짐에 따라 실시간 홍수 모니터링에 대한 수요가 증가하고 있다. 합성개구레이다는 광원과 날씨에 무관하게 지속적으로 촬영이 가능한 레이다로, 수재해가 발생하였을 때에도 영상을 제공할 수 있다. 이에 합성개구레이다를 활용한 수체 탐지 알고리즘 개발이 활발히 연구되어 왔다. 특히 딥러닝의 발달로 CNN을 활용한 수체 탐지 알고리즘이 연구됨에 따라, 높은 정확도로 수체 탐지가 기능해졌다. 하지만, CNN 기반 수체 탐지 모델은 훈련 시 높은 정량적 정확성 지표를 달성하여도 추론 후 정성적 평가 시 경계와 소하천에 대한 정확성이 떨어진다. 홍수 모니터링에서 특히 중요한 정보인 경계와 좁은 하천에 대해서 탐지의 정확성이 떨어짐에 따라 실생활 적용이 어렵다. 이에 우리는 경계를 강화한 적대적 학습 기반의 수체 탐지 모델을 개발하여 쉽게 탐지되지 않았던 부분까지 탐지하고자 한다. 적대적 학습은 생성적 적대 신경망(GAN)의 두 개의 모델인 생성자와 판별자가 서로 관여하며 더 높은 정확도를 달성할 수 있도록 학습하는 과정을 의미한다. 판별자는 생성자의 추론 결과와 실제 라벨 데이터를 구분하기 위해 학습하는 반면, 생성자는 판별자를 속이기 위해 더 실제 데이터 같은 가짜 데이터를 생성하고자 노력한다. 이러한 적대적 학습 개념을 수체 탐지 모델에 처음으로 도입하여, 생성자는 실제 라벨 데이터와 유사하게 수체 경계와 소하천까지 탐지하고자 학습한다. 반면 판별자는 경계 거리 변환 맵과 합성개구레이다 영상을 기반으로 라벨데이터와 수체 탐지 결과를 구분한다. 이때 경계 거리 변환 맵은 작은 하천과 경계에 가중치를 준 이미지로, 판별자로 하여금 판별시 작은 영역까지 고려할 수

있도록 강조하는 동시에 오탐지에 대해 억제할 수 있는 역할을 위해 제안하였다. 경계가 강조된 방향으로 적대적 학습 과정이 진행될 수 있도록, Binary Cross Entropy 손실 함수, Hausdorff distance 기반 손실 함수 그리고 적대적 손실 함수를 융합한 하이브리드 손실 함수를 새롭게 구성하였다. 제안 모델이 경계와 소하천을 정확히 탐지하는지 판단하기 위해, 정량적 지표로 F1-score, Boundary IoU, Matthews Correlation Coefficient를 사용하였으며, 육안 판독을 통해 정성적 평가도 진행하였다. 이를 통해 제안한 모델이 경계 및 소하천까지 정확하게 탐지해냄을 증명하였다. 실제 홍수 탐지에 사용하기 위해선 패치 단위 이미지가 아닌 전체 SAR 영상에서도 높은 정확도를 유지하는지 확인이 필요하다. 이를 위해 패치 단위로 학습된 모델이 전체 SAR 영상을 탐지할 수 있도록 추가 코드를 개발하여, 학습자료에 전혀 사용되지 않은 한반도를 촬영한 6개의 SAR 영상을 활용하여 탐지 결과를 비교하였다. 평가 결과 제안한 경계 강화 적대적 수체 탐지 모델이 기존 모델 대비 경계와 위양성 오류에 대해 올바르게 탐지하는 것을 증명하였다. 또한 다양한 스케일의 수체에 대해서도 꾸준히 높은 정확성을 유지하여 실제 홍수탐지를 위한 기반 모델로의 가능성을 보여주었다.

주요어: 수체 탐지, 합성개구레이다, 원격탐사, 적대적 학습, 홍수 모니터링, 딥러닝

학 번: 2021-29595