

아리랑 5호 위성 영상에서 수계의 의미론적 분할을 위한 딥러닝 모델의 비교 연구

김민지^{*}, 김승규^{**}, 이도훈^{***}, 감진규^{****}

Comparative Study of Deep Learning Model for Semantic Segmentation of Water System in SAR Images of KOMPSAT-5

Min-Ji Kim^{*}, Seung Kyu Kim^{**}, DoHoon Lee^{***}, Jin Kyu Gahm^{****}

ABSTRACT

The way to measure the extent of damage from floods and droughts is to identify changes in the extent of water systems. In order to effectively grasp this at a glance, satellite images are used. KOMPSAT-5 uses Synthetic Aperture Radar (SAR) to capture images regardless of weather conditions such as clouds and rain. In this paper, various deep learning models are applied to perform semantic segmentation of the water system in this SAR image and the performance is compared. The models used are U-net, V-Net, U2-Net, UNet 3+, PSPNet, Deeplab-V3, Deeplab-V3+ and PAN. In addition, performance comparison was performed when the data was augmented by applying elastic deformation to the existing SAR image dataset. As a result, without data augmentation, U-Net was the best with IoU of 97.25% and pixel accuracy of 98.53%. In case of data augmentation, Deeplab-V3 showed IoU of 95.15% and V-Net showed the best pixel accuracy of 96.86%.

Key words: Semantic Segmentation, Deep Learning Model, Water System, Synthetic Aperture Radar, KOMPSAT-5

1. 서 론

우리는 지구온난화를 겪으며 기후 위기의 시대를 살고 있다. 행정안전부 국립재난안전연구원이 발표한 ‘미래안전이슈(Future Safety Issue)’ 보고서에 따르면 위험성이 증가할 미래 재난 및 피해 중 1위는

홍수(18.6%)라고 한다. 그다음은 태풍(17.2%), 감염병(16.7%), 폭염(14.9%), 가뭄(11.1%) 등의 순이었다. 홍수에 따른 구체적 피해로는 도심 침수의 가능성이 가장 컸고 기반 시설 침수, 대규모 정전, 교량 및 도로 파손, 제방 붕괴 등도 언급됐다. 가뭄으로 인한 피해는 식수난 및 용수 부족과 식량 안보, 곡물

※ Corresponding Author: Jin Kyu Gahm, Address: (46241) Computer Science and Engineering Department, Pusan National University, 2 Busandaehak-ro 63 beon-gil, Geumjeong-gu, Busan, South Koea, TEL: +82-51-510-2292, FAX: +82-53-517-2431, E-mail: gahmj@pusan.ac.kr

Receipt date: Feb. 13, 2022, Approval date: Feb. 23, 2022

^{*} Major of AI, Dept. of Information Convergence Engineering, Pusan National University (Contributed equally) (E-mail: tooktooke@pusan.ac.kr)

^{**} Major of AI, Dept. of Information Convergence Engineering, Pusan National University (Contributed equally) (E-mail: aszass5387@pusan.ac.kr)

^{***} School of Computer Science and Engineering, Pusan National University
(E-mail: dohoon@pusan.ac.kr)

^{****} School of Computer Science and Engineering, Pusan National University

※ This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2020-0-01450, Artificial Intelligence Convergence Research Center [Pusan National University]).

가격 폭등 등이었다. 연구원은 과거에는 재난이 하나씩 발생했던 것에 비해 최근에는 동시에 다양한 형태로 발생해 이에 따른 예방과 대응이 필요하다고 분석했다[1].

홍수 및 가뭄으로 인한 재해의 대응 시 중요한 것 중 하나는, 강물이 어디까지 불어나거나 줄어들었는지 그 규모를 파악하는 것이다. 이를 효과적으로 파악하는 방법은 위성 영상을 이용해 수계를 한눈에 검출하는 것이다. 이때 딥러닝을 활용한 의미론적 분할이 큰 도움이 된다. 의미론적 분할은 영상의 모든 픽셀을 원하는 범주별로 분류하는 것이다. 이를 위성 영상에 적용한다면, 수계 영역과 그 외의 영역으로 픽셀들을 분류하여 신속하게 강물의 범위를 파악할 수 있다.

그러나 위성 영상에서 딥러닝 모델을 활용하여 수계를 의미론적으로 분할하는 연구는 아직 부족한 실정이다. 해외에서 연구가 조금 이루어졌지만, 국내 수계 데이터가 아닌 해외지형의 데이터를 기반으로 한 연구이고, 아리랑 5호 위성의 국내 영상 데이터를 활용한 한국항공우주연구원의 연구에서는 모델이 낮은 성과를 보였다.

본 논문에서는 아리랑 5호 위성 영상에 수계를 분할하기 위해 다양한 딥러닝 모델을 적용하여 높은 성과를 내고 이를 비교하였다. 그 방법으로 U-net, V-Net, Deeplab-V3, PAN 등 8개 모델을 사용하였다. 또한 탄성 변환을 활용하여 데이터를 증강해 데이터를 증강하지 않았을 때의 모델 성능과도 비교하였다. 결과로는 데이터 증강이 이루어지지 않고 학습된 모델 중에서는 U-Net이 97.25%의 IoU와 98.53%의 픽셀 정확도를 보이며 8가지 모델 중 가장 우수했다. 증강된 데이터에서 학습된 모델 중에서는 Deeplab-V3가 95.15%의 IoU를, V-Net이 96.86%의 픽셀 정확도를 보이며 가장 우수한 성과를 내었다.

2. 관련 연구

영상의 의미론적 분할(Semantic Segmentation)에 관한 다양한 딥러닝 모델들을 제시한 연구들은 다음과 같다. 의료 영상의 의미론적 분할 분야에서는 Olaf Ronneberger[6]가 U-net을 제안하여 당시 State-Of-The-Art(SOTA)를 달성하였다. 기존의 FCN 모델에서 벗어나 처음으로 인코더(Encoder)와

디코더(Decoder) 형태의 구조를 제안했다. 이미지를 저해상도(Low Resolution) 표현 단계로 낮추는 인코딩 과정을 거친 후, 디코딩 과정에서는 이를 고해상도(High Resolution) 규모로 확장한다. 이 과정에서 이미지의 정보가 일부는 소실될 수밖에 없으므로 이를 보완하고자 스킵 커넥션(skip connection)을 활용하였다. 인코딩 단계에서 추출한 특징 맵을 스킵 커넥션하여 같은 해상도(resolution)의 디코딩 단계에 컨캐터네이션(concatenation)하는 것이다. 권소윤[18]은 유방 영상에서 종괴의 의미론적 분할 연구를 진행하였다. Deconvolution network와 U-net 모델을 활용하였고, Deconvolution net에서 91.78%의 정확도를, U-net에서 90.09%의 정확도를 달성하였다.

Liang-Chieh Chen[2]은 Deeplab-V3를 제안하였다. 해당 모델은 Atrous Spatial Pyramid Pooling (ASPP)을 활용하여 각 컨볼루션(Convolution)을 어트루스 컨볼루션 (Atrous Convolution)으로 변경하여 수용 영역(Receptive Field)을 넓힐 수 있었다. 이러한 모델을 바탕으로 저자들은 당시 PASCAL VOC 2012 데이터셋에서 SOTA를 달성하였다. Jindong Wang[3]은 기존의 모델들이 제안한 방식인 이미지의 고해상도 표현을 저해상도 표현으로 바꾸고 이를 다시 회복시키는 과정이 의미론적 분할 결과에 좋지 않은 영향을 끼친다고 보았다. 따라서 Jindong Wang이 제안한 HRNet-V2 (High-Resolution Network)에서는 고해상도 표현을 보존시키기 위해 각 해상도에 따른 특징 맵에 대한 컨볼루션(Convolution) 연산을 병렬적으로 진행하였고 이를 통해 Cityscapes, PASCAL-CONTEXT 데이터셋 의미론적 분할 분야에서 당시 모두 SOTA를 달성하였다.

권형준[17]은 SAR 위성 영상에 조건부 생성적 적대 신경망(conditional Generative Adversarial Network, cGAN)을 적용하여 데이터 증강을 진행하여 선박을 탐지하는 연구를 진행하였다. 그러나 SAR 위성 영상에 딥러닝을 적용하여 의미론적 수계 분할을 연구한 사례는 적은 실정이다. Guojie Wang[7]은 위성 영상 수계 DenseNet 모델을 활용한 모델을 제안하였다. 이 모델은 DenseNet을 중추로 하여 특징 추출을 진행한다. 해당 모델에서는 4개의 Dense block을 활용하여 특징 추출을 진행하고 Transition block을 활용하여 각각의 Dense Block을 연결하였다. 그러나 국내가 아닌 해외의 위성 영상에 적용한

모델이다. 김명언[4]은 국내 SAR 영상을 바탕으로 U-Net, FCN-VGG16, DeepUNet 등을 활용하여 1024×1024 크기 이미지의 의미론적 수계 분할을 진행하였다. 해당 논문에서는 적은 데이터 수로 인하여 데이터를 상하좌우로 뒤집고(flip) 회전시키는(Rotation) 등의 데이터 증강을 진행하였다. 하지만, 너무 적은 원본 데이터 수로 인해 FCN-VGG16의 IoU는 67.72%, U-Net의 IoU가 65.88%로 나오는 등 성능은 확보하지 못한 것으로 보인다. 따라서 국내의 위성영상에서 수계를 의미론적으로 분할하기 위해 8개의 딥러닝 모델을 적용하여 우수한 성과를 내고 그 결과를 비교하고자 한다.

3. 연구 방법

본 논문에서는 SAR 위성 영상에서 수계를 검출하기 위해 총 8가지의 딥러닝 모델을 적용하였다. 의미론적 분할의 대표적인 모델인 U-Net과 U-Net 기반의 구조를 가진 V-Net, U²-Net, UNet 3+를 실험에 사용하였다. 또한 피라미드 구조를 활용한 의미론적 분할 모델인 PSPNet, PAN과 어트러스 컨볼루션(Atrous Convolution)을 적용한 의미론적 분할 모델인 Deeplab-V3, Deeplab-V3+로도 실험을 진행하였다. 아래는 본 논문에서 사용된 모델에 대해 간략하게 설명한 것이다.

3.1 U-Net[6] 및 UNet 3+[13]

모두 의료 영상 분야에서 의미론적 분할을 하기 위해 고안된 아키텍처들이다. U-Net은 의료 영상 분야에서 인코더-디코더 구조(Encoder-Decoder Architecture)에서 스킵 커넥션을 추가하여 좋은 성능을 얻은 모델로써 현재 의미론적 분할 분야에서 많이 사용되는 구조를 가진 모델이다. UNet 3+는 전체 규모에서 스킵 커넥션을 진행하여, 다른 규모의 특징 맵에서 높은 수준의 의미와 낮은 수준의 세부 사항을 통합한다.

3.2 V-Net[8]

V-Net은 3D 의료 영상을 의미론적 분할하기 위해 고안된 모델이다. 자기공명영상에서 전립선 검출을 위해 V 모양의 네트워크를 설계했다. 또한 Dice 계수를 활용하여 최적화를 진행하여 전경 복셀과 배경

복셀의 수 사이의 심한 불균형을 처리하였다. 또한 무작위 비선형 변환 및 히스토그램 매칭을 적용하여 데이터를 증강했다.

3.3 U²-Net[11]

U²-Net은 RSU(Residual U-block)를 통해 다양한 크기의 수용 영역이 혼합되어 다양한 크기의 정보를 받아들이는 방법을 제안하였다. RSU는 잔여 블록(Residual Block)에서 기존 하나의 컨볼루션으로 이루어져 있던 부분을 U모양의 컨볼루션으로 이루어진 블록으로 대체하고 지역 특징을 더하는 방식이다. 이를 통해 계산량을 많이 증가시키지 않으면서 더 깊은 모델을 만들 수 있었다.

3.4 PSPNet(Pyramid Scene Parsing Network)[12]

PSPNet은 주변 상황과 맞지 않는 픽셀 클래스 분류(Mismatched Relationship), 혼동되는 분류(Confusion Categories), 눈에 잘 안 띄는 객체의 클래스 분류(Inconspicuous Classes)에서 발생하는 에러를 개선하기 위해 전역 정보(Global Context)를 활용하는 피라미드 풀링 모듈(Pyramid Pooling Module)을 제안하였다. 이는 컨볼루션을 통과한 특징 맵에 각 다른 크기의 풀링을 적용하여 각기 다른 크기의 특징 맵을 만든다. 그리고 보간 방법을 통해 각 특징 맵들의 크기를 같아지게 하여 피라미드처럼 쌓는 방식이다. 기존의 FCN(Fully Convolutional Networks for Semantic Segmentation)에 비해 높은 정확도를 달성하지만, 상대적으로 느린 수행 속도를 가지는 점이 단점이다.

3.5 Deeplab-V3[2] 및 Deeplab-V3+[10]

DeeplabV3는 기존 컨볼루션 대신 어트러스 컨볼루션 방법을 제안하였다. 해당 논문에서 말하는 어트러스 컨볼루션이란 기존 컨볼루션에서 진행할 때는 커널(kernel)과 인풋 이미지의 위치가 같은 곳에 대해 1대1 대응으로 연산하였지만, 연산되는 이미지의 픽셀 간격을 띄워 모델이 더 넓은 수용 영역(Receptive Field)을 반영할 수 있도록 개선한 것이다. 또한, 인코더 단에서 ResNet을 중추로 하여 좀 더 조밀(Dense)한 특징 맵(Feature Map)을 얻을 수 있었다. Deeplab-3+에서는 깊이 컨볼루션(Depthwise Con-

volution)과 어트러스 컨볼루션을 결합한 어트러스 세파러블 컨볼루션 (Atrous Separable Convolution)을 제안하여 성능을 유지하면서 연산량을 줄일 수 있었다.

3.6 PAN(Pyramid Attention Network)[9]

PAN 또한, PSPNet과 마찬가지로 전역 정보를 활용하기 위한 특징 피라미드 어텐션 (Feature Pyramid Attention)을 제안하였다. 이 모델은 어텐션을 통해 픽셀 레벨에서의 정보를 얻고 피라미드 구조를 활용하여 수용 영역을 늘려 좋은 성능을 달성하였다.

4. 실험 및 연구 결과

4.1 실험 방법

아리랑 5호 위성은 한국항공우주연구원(Korea Aerospace Research Institute, KARI)에서 2013년 8월 22일 발사된 관측 위성 다목적실용위성 5호이며 KOMPSAT-5 또는 K5라고 부른다. 아리랑 5호는 아리랑 1호, 2호와 달리 SAR 센서를 이용해 영상을 생성한다. SAR는 Synthetic Aperture Radar의 준말로 공중에서 지상 및 해양을 관찰하는 합성 개구 레이더이다. SAR는 전자기파를 이용하는 능동 센서로 가시광선 보다 파장이 긴 마이크로파를 이용하기 때문에 광학 센서와는 다르게 비, 구름, 안개 등 기상 조건이나 일조 현상과 관계없이 전천후로 지역의 영상 획득이 가능하다.

아리랑 5호의 SAR 영상은 AI Hub(<https://aihub.or.kr/>)에서 제공하는 한국항공우주연구원이 구축한 '위성영상 객체 판독 데이터셋'에서 획득했다. 이 데이터셋은 아리랑 5호의 SAR 영상으로부터 얻은 1024×1024 크기의 패치들이다. 패치는 도심, 농경지, 초지, 산지 등의 다양한 지역에서 추출되었고, 레이블 이미지의 경우 0과 1로 이루어진 이진 영상(Binary Image)이다. 데이터 셋은 총 2,701장으로 2,401장의 학습 데이터와 300장의 테스트 데이터가 있다.

4.1.1 데이터 증강-탄성 변형(Elastic Deformation)

일반적으로 딥러닝 모델에서는 데이터의 수가 늘수록 학습 성능이 올라가는 경향성을 보인다. 특히 학습 데이터 수가 적을 때 데이터 증강을 적용한다면 모델의 강건성(Robustness)을 향상하는 데에 도움

을 줄 수 있다. 따라서 더 좋은 결과를 얻기 위해 데이터 증강을 진행하였다. 이를 위해 U-Net 모델에서 저자가 세포 이미지의 의미론적 분할 학습 시 성능향상을 위해 사용한 데이터 증강 기법인 탄성 변형(Elastic Deformation) 기법을 적용했다.

그 이유는 강물 역시 시간의 흐름에 따라 그 물줄기가 조금씩 바뀌기 때문이다. 탄성 변형은 이를 자연스럽게 나타내는 방법이고, 시그마(Sigma)값을 적절히 조정하면 실제와 유사하게 물이 움직이는 것을 표현할 수 있다. Δ_x , Δ_y 는 이미지 변경 시 원본으로부터 변경되는 거리를 의미한다. 원본 이미지 I 가 주어질 때 변환된 이미지 I_{trans} 는 식(3)과 같다.

$$\Delta_x = G(\sigma) * \alpha \times Rand(n, m) \tag{1}$$

$$\Delta_y = G(\sigma) * \alpha \times Rand(n, m) \tag{2}$$

$$I_{trans}(j + \Delta_x(j, k), k + \Delta_y(j, k)) = I(j, k) \tag{3}$$

이때, $Rand(n, m)$ 은 랜덤하게 선택된 변경될 값이고 n 과 m 은 이미지의 크기다. G 는 가우시안 필터(Gaussian Filter), α 는 초기 무작위 변수 중 최댓값을 의미한다. σ 값을 조절하여 평활화 작업의 강도를 설정할 수 있다. j, k 는 이미지의 픽셀 위치이다[16]. 본 실험에서 사용한 시그마 값은 3, 그리드 값은 3×3으로 설정하였으며 이때 사용된 모드는 반사(Reflection)를 사용하였다. 반사 모드는 탄성 변형 시 발생하는 가장자리의 빈 공간을 채우기 위해 사용되었으며 빈 공간과 대칭되는 부분을 반사하여 복사하는 방식을 말한다.

Fig. 1는 데이터 증강이 이루어지는 것을 시각적으로 확인할 수 있다. 이러한 방식을 통해 기존의 학습 데이터를 3배로 늘렸다. 탄성 변형 자체가 무작위로 작동하는 방식이므로 원본, 탄성 변형 이미지-1, 탄성 변형 이미지-2로 하여 총 7,203장의 학습 데이터를 얻을 수 있었다. 또한, 테스트셋에는 변형을 적용하지 않고 실제 이미지를 바탕으로 결과를 도출할 수 있도록 설계하였다. 탄성 변형을 각 이미지에 적용하였을 때 발생하는 변화에 대한 추가적인 예시는 Fig. 2를 참조하길 바란다.

4.2 평가

의미론적 분할의 결과를 비교하기 위해서는 얼마나 레이블과 유사한지에 대한 지표가 필요하다. 본 논문은 널리 쓰이는 평가지표인 IoU (Intersection

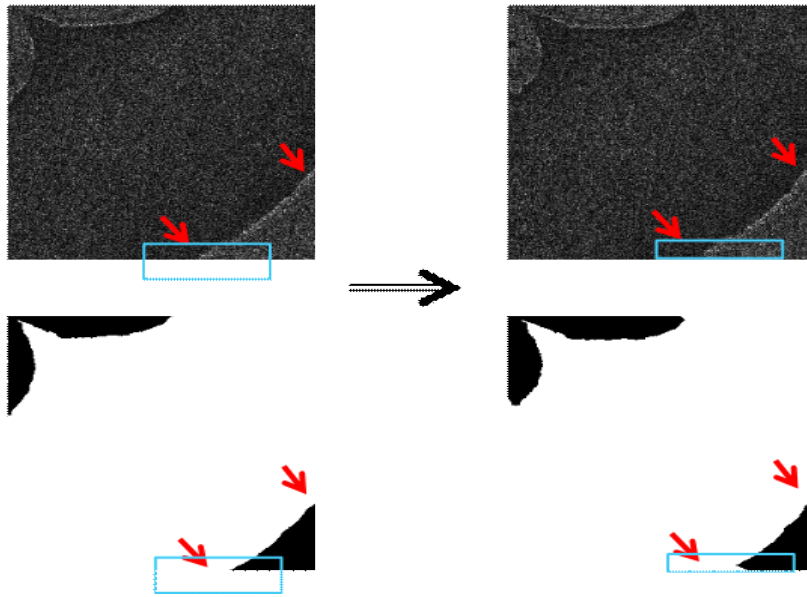


Fig. 1. Results of Elastic Deformation.

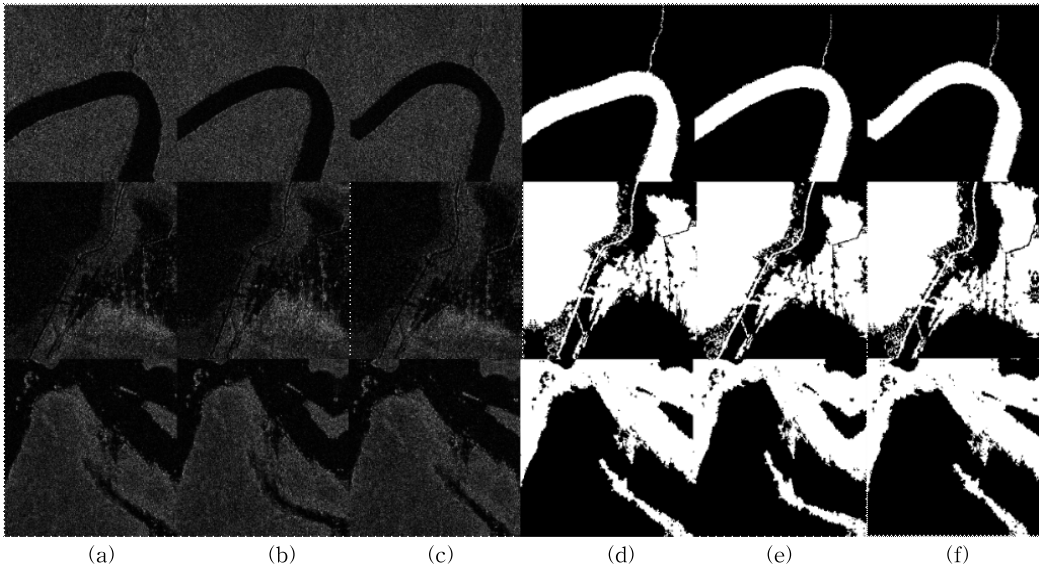


Fig. 2. Results of Elastic Deformation, (a), (d) are the original images and its labels, (b), (e), are the first augmented images, (c), (f) are the second augmented images, (Sigma=3, Grid=3x3)

over Union)와 픽셀 정확도를 평가지표로 선택하였다. IoU란 두 개의 이미지가 있을 때 두 이미지의 교집합과 합집합의 비로 두 이미지가 얼마나 서로 겹치는지를 정량적으로 표시할 수 있다. 이는 픽셀별로 진행되며, 결과가 1일 때 완전히 같은 이미지라고 볼 수 있다. 두 이미지 A와 B의 겹침 정도 IoU(A,

B)는 식 (6)과 같이 나타낸다.

$$IoU(A,B) = \frac{A \cap B}{A \cup B} \quad (6)$$

픽셀 정확도 PixelAccuracy는 두 이미지의 각 위치에 대한 픽셀값이 같은지에 대한 지표이다. 이는 레이블 전체 픽셀 개수에 대한 레이블과 같은 픽셀

개수의 비로써 식 (7)과 같다.

$$\text{PixelAccuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (7)$$

4.3 실험 결과

본 논문은 U-net, V-Net, U²-Net, UNet 3+, PSPNet, Deeplab-V3, Deeplab-V3+, PAN 총 8개의 모델을 활용하여 실험을 진행하였다. 실험을 진행하기 위해 기존 1024×1024 크기의 이미지를 256×256으로 변경하였고, 배치 크기는 8, 학습률(Learning rate)은 0.0001, 최적화 함수는 Adam[14], 손실함수는 분할(Segmentation)을 위한 IoU(Intersection over Union)[15]를 손실함수로 사용하여 각 모델에 대해 동일한 환경에서 실험을 진행하였다. 그리고 앞선 연구와의 비교를 위해 김명언[4]의 논문의 결과를 Table 1에 나타내었다.

U-net 모델로 학습을 진행하였을 때 데이터 증강 전과 후의 이미지를 Fig. 3에서 확인할 수 있다. 하지만 큰 물줄기와 따로 떨어진 수계에서는 분할이 잘 진행되지 않은 것을 확인할 수 있었다. 모델별 실험 결과에 대한 IoU 값과 픽셀 정확도 값은 Table 1에 나타내었다. 데이터 증강이 이루어지지 않은 데이터 셋으로 학습시켰을 때, 모델 모두 95% 이상의 높은 IoU(Intersection of Union)를 달성하였으며 특히 U-net은 97.25%로 가장 높은 IoU를 달성했다. 픽셀 정확도 역시 모든 모델이 91% 이상을 달성했으며, U-

net이 98.53%로 가장 높은 픽셀 정확도를 기록했다.

반면 데이터 증강이 이루어진 데이터 셋으로 학습시켰을 때 모델의 IoU는 약 평균 0.048, 픽셀별 정확도는 평균 0.025만큼 소폭 감소했다. 이때 가장 높은 IoU를 달성한 모델은 Deeplab-V3로 95.15%였고, 가장 높은 픽셀 정확도를 달성한 모델은 V-net으로 96.86%를 달성했다.

데이터 증강 시에 모델 성능이 소폭 하락하였다는 것은 일반화(Generalization) 성능이 저하되었다는 것을 의미한다. Olaf Ronneberger[6]는 세포 이미지 데이터셋을 증강하기 위해 탄성 변환을 적용했고 시그마를 10으로 설정하여 성공적으로 데이터를 증강했다. 본 연구에서도 탄성 변환 시 시그마를 10으로 하여 SAR 영상에 적용해보았으나, 현실 수계와는 매우 다른 어색한 이미지가 생성되어 최대한 자연스러운 이미지를 생성하도록 시그마를 3으로 적용한 것이었다. 그러나 Fig. 2에서 확인할 수 있는 것처럼 원본 이미지와 증강된 이미지가 큰 차이를 보이지 않았고 오히려 이것이 증강된 데이터로 학습된 모델의 성능 저하에 영향을 미친 것으로 보인다. 따라서 후속 연구에서는 시그마값을 3이 아닌 다른 값으로 설정하거나 탄성 변환이 아닌 다른 데이터 증강 방법을 적용해 볼 수 있을 것이다.

5. 결 론

본 논문은 홍수 또는 가뭄과 같은 자연재해에 대

Table 1. The performance for Various Deep Learning Models.

| | Model | Not Augmented | | Augmented | |
|--------|-------------------------|---------------|---------------|---------------|---------------|
| | | IoU | Pixel Acc. | IoU | Pixel Acc. |
| 김명언[4] | FCN-VGG16 | - | - | 0.6772 | - |
| | FCN-VGG19 | - | - | 0.6679 | - |
| | U-Net | - | - | 0.6588 | - |
| | DeepUNet | - | - | 0.6469 | - |
| Ours | U-Net[6] | 0.9725 | 0.9853 | 0.9351 | 0.9655 |
| | V-Net[8] | 0.9702 | 0.9840 | 0.9417 | 0.9686 |
| | U ² -Net[11] | 0.9692 | 0.9837 | 0.8969 | 0.9368 |
| | UNet 3+[13] | 0.9659 | 0.9818 | 0.9225 | 0.9575 |
| | PSPNet[12] | 0.9700 | 0.9298 | 0.9029 | 0.8032 |
| | Deeplab-V3[2] | 0.9681 | 0.9248 | 0.9515 | 0.8987 |
| | Deeplab-V3+[10] | 0.9595 | 0.9106 | 0.8825 | 0.9463 |
| | PAN[9] | 0.9708 | 0.9324 | 0.9271 | 0.8577 |

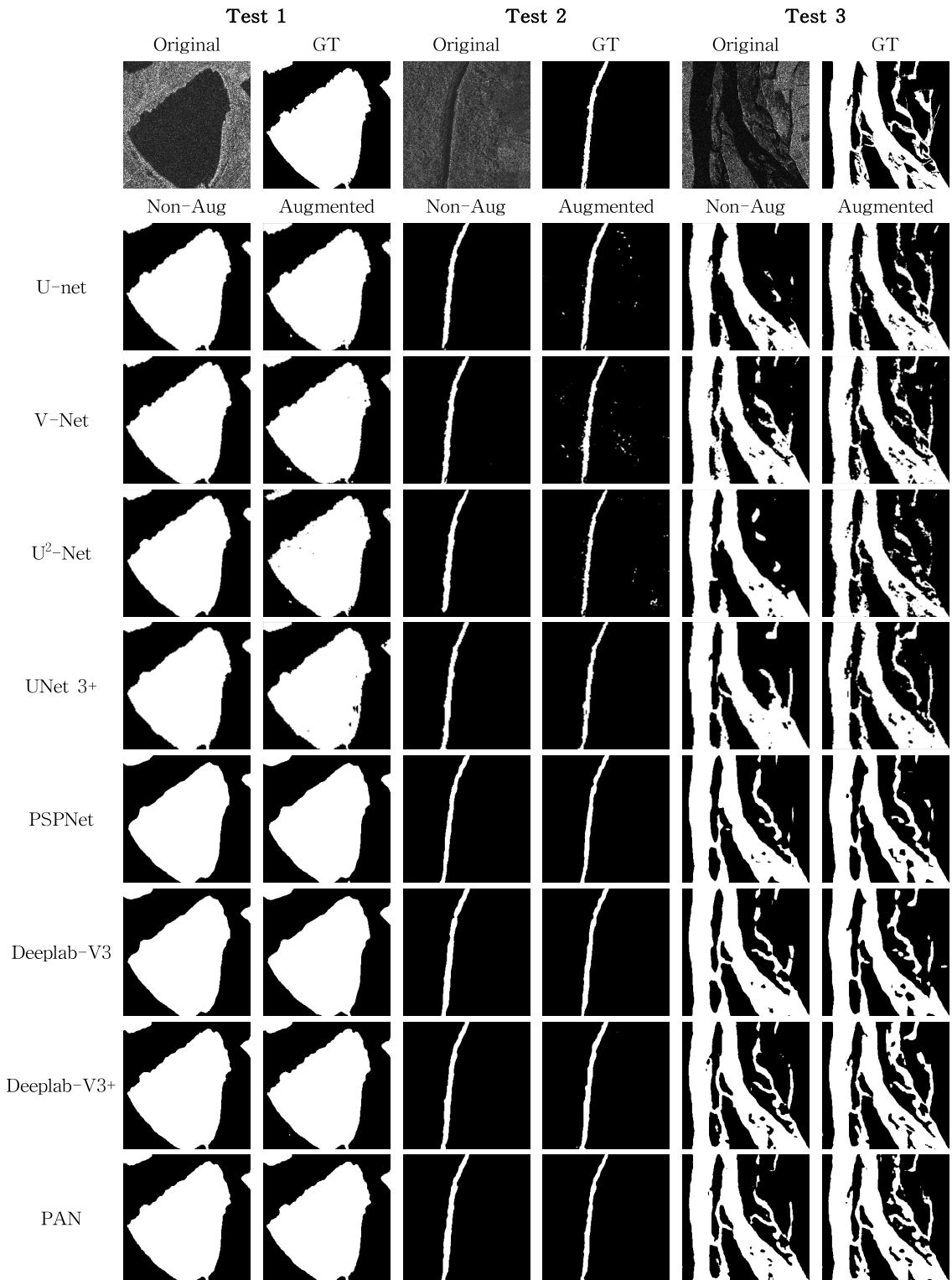


Fig. 3. Results of Water Segmentation of 8 models.

용하기 위해 SAR 위성 영상에 딥러닝을 적용하여 수계 영역을 분할하는 연구를 진행하였다. 연구에 사용된 의미론적 분할의 딥러닝 모델은 U-net, V-Net, U²-Net, UNet 3+, PSPNet, Deeplab-V3, Deeplab-V3+, PAN 이며, 탄성 변형을 적용한 데이터 증강을 통해 데이터 쌍을 3배 증가하여 각 모델의 학습 결과를 확인했다. 데이터 증강이 이루어지지 않고 학습된 모델 중에서는 U-Net이 97.25%의 IoU와 98.53%의 픽셀 정확도를 보이며 가장 우수했다. 증강된 데이터에서 학습된 모델 중에서는 Deeplab-V3가 95.15%의 IoU를, V-Net이 96.86%의 픽셀 정확도를 보이며 우수한 성과를 내었다. 본 논문이 가지는 의의는 다음과 같다.

첫째로, 13가지의 의미론적 분할의 딥러닝 모델들을 적용하여 각 모델의 결과를 비교하여 어떤 모델이 SAR 위성 영상 수계 검출에 적합한지 실험하였다. 이는 수계의 의미론적 분할을 연구하는 데 참고 자료로 사용될 수 있을 것이다. 둘째로, 자연스러운 수계의 변형을 나타내고자 탄성 변형을 활용한 데이터 증강을 통해 실험에 활용된 13가지의 모든 모델에 대해 실험을 진행하였으나 큰 성능향상을 기대할 수 없었다. 이를 통해, 해당 데이터 증강 기법이 현실을 잘 반영하지 못했다는 것을 알 수 있다.

그러므로 본 논문이 가지는 한계점은 다음과 같고 이를 향후 연구에 반영할 수 있다.

첫째로, 실제 위성 사진을 다운샘플링 하여 사용하였다. 기존의 높은 이미지 해상도를 인위적으로 낮췄으므로 모델의 성능에 영향을 미쳤을 것이다. 둘째로, 모델들이 폭이 좁은 물길은 잘 검출해내지 못하였다. 실선과 같이 얇은 물길도 잘 검출해내기 위해서는 다른 손실함수를 적용하거나 얇은 물길에 해당하는 부분에는 가중치를 더 주어 검출이 잘 되도록 할 수도 있을 것이다. 셋째로, 데이터 증강 방법을 다양하게 적용해볼 수 있을 것이다. 본 논문에서는 탄성 변형을 적용한 데이터 증강을 통해 눈에 띄는 성능향상은 없었지만 변환 조건을 다르게 하거나 다른 방법을 적용한다면 성능을 향상할 수 있을 것이다.

REFERENCE

- [1] NDMI, "Prospects of Future Disasters in Our Country through Scenarios," *Future Safety Issue*, Vol. 18, pp. 1-41, 2021.
- [2] L. Chen, G. Papandreou, F. Schroff, and H. Ada, "Rethinking Atrous Convolution for Semantic Image Segmentation," *arXiv Preprint*, arXiv:1706.05587, 2017.
- [3] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, and Y. Zhao et al., "Deep High-Resolution Representation Learning for Visual Recognition," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 43, No. 10, pp. 3349-3364, 2021.
- [4] M.Y. Kim, C.Y. Choi, S.H. Han, S.J. Lee, and O. Han, "Deep Learning based Water Segmentation of KOMPSAT-5 SAR Images," *The Korean Society for Aeronautical and Space Sciences*, pp. 131-132, 2020.
- [5] H. Lee, H. Yoon, and H. Han, "A Study on the Determination of Indicator for the Risk Assessment of Ground Depression Using SAR Image," *Journal of the Korean Geo-Environmental Society*, Vol. 22, No. 7, pp. 13-20, 2021.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, 2015.
- [7] G. Wang, M. Wu, X. Wei, and H. Song, "Water Identification from High-Resolution Remote Sensing Images Based on Multidimensional Densely Connected Convolutional Neural Networks," *Remote Sensing*, Vol. 12, No. 5, pp. 795, 2020.
- [8] F. Milletari, N. Navab, and S.A. Ahmadi. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," *Fourth International Conference on 3D Vision*, pp. 565-571, 2016.
- [9] H. Li, P. Xiong, J. An, and L. Wang, "Pyramid Attention Network for Semantic Segmentation," *arXiv Preprint*, arXiv:1805. 10180, 2018.
- [10] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-Decoder with Atrous

Separable Convolution for Semantic Image Segmentation,” *Proceedings of the European Conference on Computer Vision*, pp 801–818, 2018.

- [11] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, “U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection,” *Pattern Recognition*, Vol 106, p. 107404, 2020.
- [12] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2881–2890, 2017.
- [13] H. Huimin, L. Lin, R. Tong, H. Hu, Q. Zhang, and Y. Iwamoto et al., “Unet 3+: A Full-Scale Connected Unet for Medical Image Segmentation,” *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1055–1059, 2020.
- [14] D.P. Kingma and J.L. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv Preprint*, arXiv:1412.6980, 2014.
- [15] M.A. Rahman and Y. Wang, “Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation,” *International Symposium on Visual Computing*, pp. 234–244, 2016.
- [16] E. Castro, J.S. Cardoso, and J.C. Pereira, “Elastic Deformations for Data Augmentation in Breast Cancer Mass Detection,” *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 230–234, 2018.
- [17] H. Kwon, S. Jeong, S. Kim, J. Lee, and K. Sohm, “Deep-Learning based SAR Ship Detection with Generative Data Augmentation,” *Journal of Korea Multimedia Society*, Vol. 25, No. 1, pp. 1–9, 2022.
- [18] S. Kwon, Y. Kim, and G. Kim, “An Automatic Breast Mass Segmentation based on Deep

Learning on Mammogram,” *Journal of Korea Multimedia Society*, Vol. 21, No. 12, pp. 1363–1369, 2018.



김민지

2021년 8월 부산대학교 경영학과 학사
 2021년 9월~현재 부산대학교 정보융합공학과 석사과정
 관심분야: 컴퓨터 비전, 인공지능



김승규

2019년 부산대학교 기계공학부 학사 졸업
 2021년~현재 부산대학교 정보융합공학과 석사과정
 관심분야: 컴퓨터비전, 의료영상, 인공지능



이도훈

1986년 부산대학교 계산통계학과 학사
 1992년 부산대학교 전자계산학과 석사
 1998년 부산대학교 전자계산학과 박사

1995년~2005년 밀양대학교 컴퓨터공학과 교수
 2004년~2005년 Indiana University 방문연구원
 2006년~현재 부산대학교 정보컴퓨터공학부 교수
 관심분야: 컴퓨터비전, 데이터 분석



김진규

2000년 서울대학교 컴퓨터공학과 학사
 2014년 University of California, Los Angeles (UCLA), Computer Science 박사
 2019년~현재 부산대학교 정보컴퓨터공학부 조교수

관심분야: 컴퓨터비전, 의료영상, 인공지능