

SAR 선박 탐지를 위한 의미론적 분할 이중 디코더 모델

정진원

신요안°

승실대학교 전자정보공학부

Semantic Segmentation Double Decoder Model for SAR Ship Detection

Jung-Jin Won

Yoan Shin°

School of Electronic Engineering

Soongsil University

jinwonj@soongsil.ac.kr

yashin@ssu.ac.kr

(°Corresponding Author)

요 약

본 논문에서는 Synthetic Aperture Radar (SAR) 선박 영상의 의미론적 분할의 성능을 향상시키기 위한 새로운 모델 방식을 제안한다. 기존의 전통적인 인코더-디코더 구조를 갖고있는 모델과 다르게 2 개의 디코더를 사용하며 디코더를 연결한 구조를 만들어 특징맵에 포함된 공간 정보의 손실을 개선한 이중 디코더 모델이다. 이중 디코더 모델은 특징 학습 과정에서의 특징 손실을 줄이며 클래스 불균형 문제로 인해 객체의 분할이 잘못되거나 실패하는 문제를 개선하였다. 실험 결과, 제안하는 이중 디코더 모델이 다른 대표적 의미론적 분할 모델보다 더 좋은 탐지 성능을 보여준다.

1. 서론

Synthetic Aperture Radar (SAR)는 전천후에 구름과 밤낮의 영향을 받지 않으며, 다양한 정보 획득 능력을 갖추고 있어 군과 민간 모두에 광범위하게 연구되고 있다. 하지만 SAR 영상은 전송, 수신 및 처리 과정에서 생기는 노이즈로 인해 관측과 판별이 어렵다. 특히 SAR 선박 영상의 경우, 의미론적 분할을 적용할 때 배경과 선박 이렇게 단 2 개의 클래스만을 가지고 있기 때문에 클래스 불균형 문제가 심하고 의미론적 분할을 적용하기 어렵다. 의미론적 분할은 특정 범주의 객체 또는 영역에 따라 픽셀 수준의 단위로 클래스를 분류하는 작업을 말한다. 최근에는 딥러닝 기술의 발달로 의미론적 분할의 성능이 크게 향상되었으며, 일반적으로 전이 학습을 통해 특징을 추출하는 인코더 부분과 업샘플링을 활용해 원래 이미지 사이즈로 맞추는 디코더를 사용하는 방법들이 많이 연구되고 있다 [1].

U-Net은 인코더-디코더 구조를 따르는 가장 일반적인 의미론적 분할 모델이다 [2]. U-Net 모델은 컨볼루션 레이어 2 개와 풀링 레이어 1 개로 이루어진 인코더와 Transposed 컨볼루션 레이어 1 개와 컨볼루션 레이어 2 개 구성되어 있는 디코더를 사용해 U자 형태의 네트워크로 이루어져 있다. 인코더는 이미지의 깊이를 증가시키면서 이미지의 크기를 점진적으로 축소함으로써 이미지의 특징 정보를 추출하며 디코더는 이미지를 업샘플링함으로써 이 정보를 기

존 픽셀 위치로 다시 변환한다. 이때 깊이를 줄이면서 이미지의 크기를 증가시키며 인코더의 특징맵을 디코더에 연결하는 스킵 연결이 사용된다.

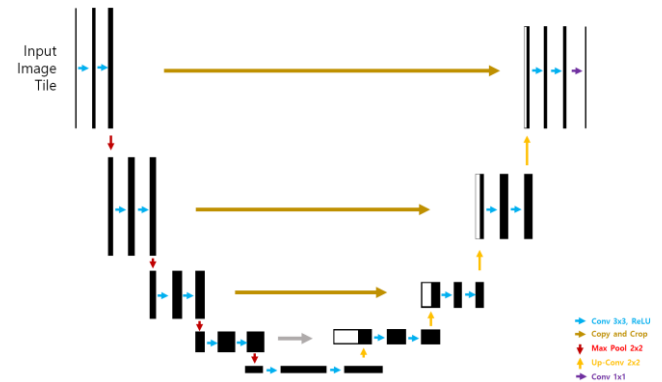


그림 1. U-Net [2]

2. 이중 디코더 모델

본 논문에서는 SAR 선박 탐지를 위한 의미론적 분할 이중 디코더 모델을 제안한다. 그림 2는 제안하는 이중 디코더 모델로 RestNet152V2 [3]층을 각각의 인코더로 사용하며 업샘플링을 개선하기 위해 인코더의 특징맵을 디코더에 연결하는 스킵 연결이 사용된다. 이 모델은 U-Net 구조와 같이 기존의 전통적인 인코더-디코더 구조를 갖고 있는 모델과 다르게 2 개의 디코더를 사용하여 이중 디코더 구조를

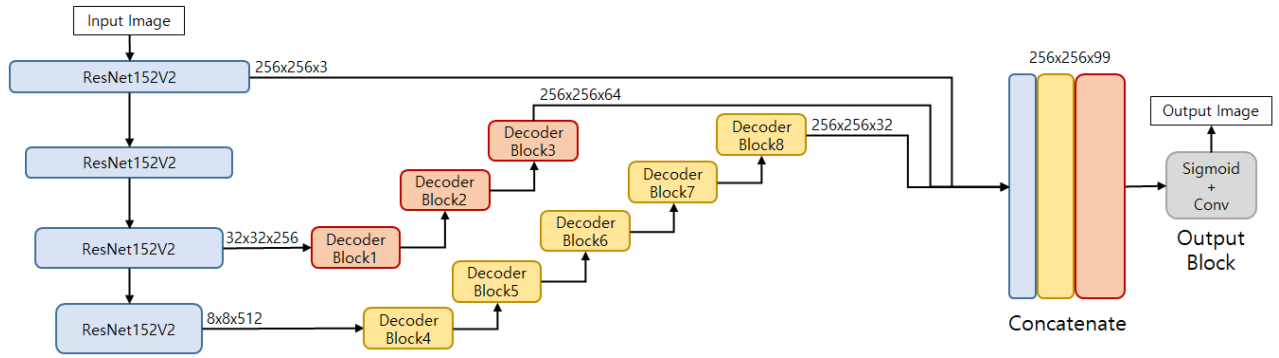


그림 2. Double Decoder Model 구조

가지고 있는 것이 특징이다. 2 개의 디코더는 서로 이어 붙여 사용하며 특징맵에 포함된 공간 정보의 손실을 줄이게 된다. 특히 특징 추출 과정에서 ResNet152V2의 잔차학습을 사용하여 특징의 손실을 줄이고 focal loss 를 사용해 클래스 불균형 문제로 인해 객체의 분할이 잘못되거나 실패하는 문제를 개선하였다. 마지막 Output Block에서는 컨볼루션 레이어를 거쳐 Output Image 를 출력한다.

그림 3 은 Decoder Block 의 구조를 나타낸다. x 는 Decoder block 으로 들어오는 특징맵으로 Transposed 컨볼루션, 3x3 컨볼루션, Batch normalization (BN), Rectified Linear Unit (ReLU)를 지난다. x 는 Transposed 컨볼루션 레이어를 통해 업샘플링되며 필터는 작아지고 특징맵이 확장된다. 그 후 3x3 컨볼루션, Batch normalization (BN), Rectified Linear Unit (ReLU)구조를 2 번 거치면서 특징을 학습한다. 이러한 과정은 이중 구조로 나누어 진행되며 최종적으로 Concatenate 를 통해 모두 연결되어 연속적인 특징 추출 및 압축 과정에서 발생하는 특징 손실과 네트워크가 깊어질수록 발생하기 쉬운 기울기 소실 문제를 최소화한다.

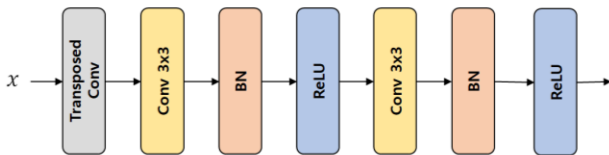


그림 3. Decoder Block 구조

3. 실험 결과 및 결론

실험은 ResNet 을 백본으로 가지는 U-Net 방식과 PSPNet 방식, FPN 방식 그리고 제안방법을 각각 비교하였다. 사용한 데이터 세트는 각종 위성으로 찍은 1-채널 회색조 SAR 선박 영상으로 구성되어 있으며, 256x256 의 크기로 총 5,604 개의 영상을 전처리한 데이터 세트로 사용하였다 [4]. 표 1 에서 볼 수 있듯이 제안하는 방법이 99.9%의 높은 정확도를 보였다. 특히, 제안하는 방법은 다른 비교 모델에 비해 선박 IoU 값이 77.3%로 선박을 보다 더 잘 식별하였으며, 비교 모델 중 가장 높은 U-Net 과 비교했을 때 약 9% 향상되었음을 알 수 있다.

본 논문에서는 SAR 선박 영상에서 제안하는 이중 디코더 모델을 활용하여 의미론적 분할의 성능 향상이 있는지를 확인하였다. 이중 디코더 모델은 잔차학습과 이중 디코더 구조를 사용하여 특징 추출 및 압축 과정에서 발생하는 특징 손실과 클래스 불균형 문제를 개선하였다. 이러한 방법을 통해 제안하는 이중 디코더 모델이 다른 의미론적 분할 모델과 비교했을 때 우수한 성능을 얻을 수 있음을 확인하였다.

표 1. 실험 결과

Scheme	Metrics		
	Accuracy	IoU (ship)	Loss
Proposed	99.9%	77.3%	0.0311
U-Net	99.8%	68.8%	0.0030
PSPNet	99.7%	42.1%	0.0003
FPN	99.8%	66.0%	0.0001

Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기 획평가원의 대학 ICT 연구센터육성지원사업의 연구 결과로 수행되었음 (IITP-2023-2018-0-01424)

References

- [1] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," Artificial Intell. Rev., vol. 52, pp. 1089-1106, Aug. 2019.
- [2] D. Raychaudhuri, "Performance analysis of random access packet-switched code division multiple access systems," IEEE Trans. Commun., vol. COM-29, pp. 895-901, Jun. 1981.
- [3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE CVPR 2016, pp. 770-778, Las Vegas, USA, June. 2016.
- [4] S. Wei, X. Zeng, Q. Qu, M. Wang, H. Su, and J. Shi, "HRSID: A high-resolution SAR images dataset for ship detection and instance segmentation," IEEE Access, vol. 8, pp. 120234-120254, June. 2020.