다중 편광 SAR 영상 목표물 인식을 위한 딥 컨볼루션 뉴럴 네트워크

염광영, 김문철
한국과학기술원 전기 및 전자 공학과
yky5464@kaist.ac.kr, mkim@ee.kaist.ac.kr

Convolutional neural network for multi polarization SAR recognition

Gwang-Young Youm, Munchurl Kim
Korea Advanced Institute of Science and Technology School of Electrical Engineering

요약

최근 Convolutional neural network (CNN)을 도입하여, SAR 영상의 목표물 인식 알고리즘이 높은 성능을 보여주었다. SAR 영상은 4 종류의 polarization 정보로 구성되어있다. 기계와 신호처리의 비용으로 인하여 일부 데이터는 적은 수의 polarization 정보를 가지고 있다. 따라서 우리는 SAR 영상 data 를 멀티모달 데이터로 해석하였다. 그리고 우리는 이러한 멀티모달 데이터에 잘 작동할 수 있는 콘볼루션 신경망을 제안하였다. 우리는 데이터가 포함하는 모달의 수에 반비례 하도록 scale factor 구성하고 이를 입력 크기조절에 사용하였다. 입력의 크기를 조절하여, 네트워크는 특징맵의 크기를 모달의 수와 상관없이 일정하게 유지할 수 있었다. 또한 제안하는 입력 크기조절 방법은 네트워크의 dead filter 의 수를 감소 시켰고, 이는 네트워크가 자신의 capacity 를 잘 활용한다는 것을 의미한다. 또 제안된 네트워크는 특징맵을 구성할 때 다양한 모달을 활용하였고, 이는 네트워크가 모달간의 상관관계를 학습했다는 것을 의미한다. 그 결과, 제안된 네트워크의 성능은 입력 크기조절이 없는 일반적인 네트워크보다 높은 성능을 보여주었다. 또한 우리는 전이학습의 개념을 이용하여 네트워크를 모달의 수가 많은 데이터부터 차례대로 학습시켰다. 전이학습을 통하여 네트워크가 학습되었을 때, 제안된 네트워크는 특정 모달의 조합 경우만을 위해 학습된 네트워크보다 높은 성능을 보여준다.

1. 서론

SAR (Synthetic Aperture Radar) 는 날씨와 밤낮에 관계 없이 사용 가능하고 뛰어난 관통력을 가지고 있어 국방분야에서 적을 감시하는데 널리 사용되고 있다. 하지만 SAR 영상은 심한 Speckle 잡음으로 인하여, 직관적인 해석이 어렵다. 따라서 SAR 영상을 해석하고 분석하는 알고리즘에 관한 수요가 발생하게 되었다.

최근 딥 러닝 기술이 다양한 물체 인지 및 인식 문제에 적용되어 괄목할만한 성능을 보여주었다 [1, 2]. 그리고 SAR 영상을 식별 문제에 딥 러닝 기술을 사용한 알고리즘 역시 연구되어 왔다 [3, 4]. Ding 은 딥러닝 기술을 SAR 영상 인식 문제를 적용할 수 있도록 SAR 영상 데이터 증대 알고리즘을 고안하였다 [3]. 또한 Chen 은 대부분의 학습 파라미터가 fully connected 층에서 발생하는 점에 착안하여, 모든 층을 convolution 층으로 구성한 네트워크를 고안하였다 [4].

SAR 영상은 크게 4 가지 polarization 데이터로 이루어져있다: HH, HV, VH, VV [5]. 하지만 모든 polarization 데이터를 얻는 것은 기계와 신호처리의 비용으로 인하여 쉽지 않다. 또한 전시상황에서는 기계와 통신장비의 결함으로 인하여 일부의 polarization 데이터가 손실되는 경우도 발생할 수 있다. 하지만 기존의 SAR 영상 식별 연구들은 1 가지의 polarization 를 처리한다.

따라서 본 논문에서는 SAR 영상을 multi modal data 로 해석하고, 이러한 문제 어떠한 polarization 데이터의 조합에도 강인하게 처리할 수 있는 네트워크를 고안하였다. 또한, 본 논문에서는 일반적인 convolution neural network (CNN)의 특징맵이 존재하는 모달에 수에 따라서 크기가 달라질 수 밖에 없다는 것을 확인하였다. 이것이 네트워크의 성능저하를 발생시킨다고 가설을 세우고, 입력 크기조절 네트워크를 고안하여, 이러한 문제를 해결하였다. 제안된 네트워크는 일부의 polarization 정보만 존재할 때도 잘 작동할 수 있으며, 여러 polarization 가 존재한다면 이러한 다중 정보를 잘 활용하여 SAR 영상을 식별할 수 있다. 또한 전이학습을 통하여 제안된 네트워크를 학습시켰을 때, 네트워크는 특정 모달의 경우만을 위해 학습된 네트워크보다 높은 성능을 보여주었다.

2. 제안하는 방법

본 논문은 SAR 영상이 일부의 polarization 정보가 부재 할 수 있다는 사실을 바탕으로 이를 멀티모달 데이터로 해석하였다. 그리고 이러한 멀티모달 데이터를 잘 처리할 수 있는 알고리즘을 고안하였다. 그림 1 은 본 논문에서 사용한 네트워크 구조를 보여준다. 우리는 입력을 4 channel 로 구성하여 네트워크가 멀티 polarization 정보를 활용하도록 하였다. 그리고 존재하지 않는 polarization 데이터는 0 으로 대체하였다.

우리는 먼저 성능의 비교를 위하여 네트워크를 특정 타입의 polarization 조합 만을 위하여 작동하는 네트워크를 학습시켰다: HH, HV, VH, VV, All. 여기서 All 은 4 개의 polarization 데이터가 모두 존재하는 경우이다. 표 1 은 특정 polarization 조합만을 처리하는 네트워크의 성능을 보여준다. 우리는 이 성능을 성능의 지표로 삼고 실험을 진행하였다.

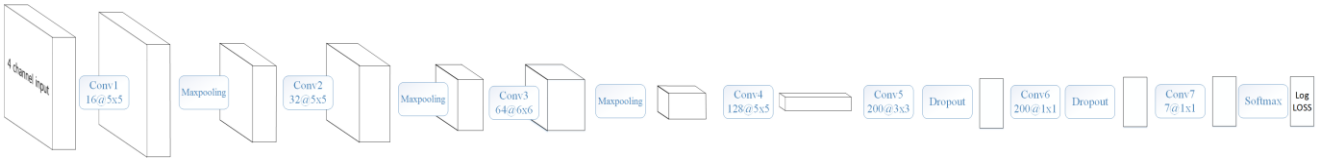
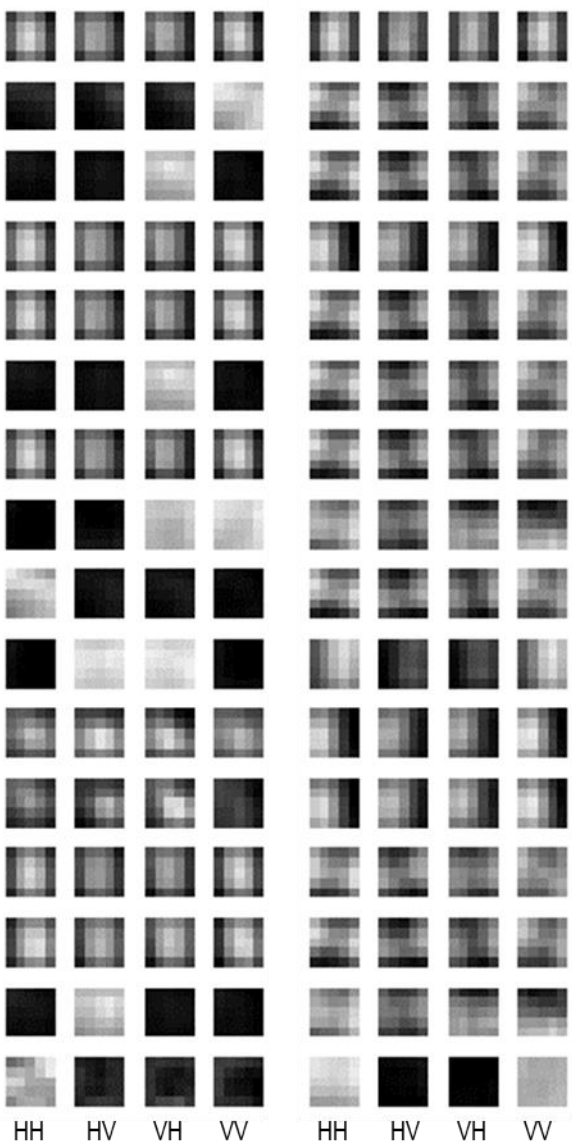


그림 1. SAR 영상 인식을 위한 네트워크 구조



(가) (나)

그림 2. 학습된 네트워크의 weight (가): 입력 크기조절이 없는 네트워크 (나) 입력 크기조절 네트워크

존재하는 경우이다. 표 1 은 특정 polarization 조합만을 처리하는 네트워크의 성능을 보여준다. 우리는 이 성능을 성능의 지표로 삼고 실험을 진행하였다.

그리고 우리는 모든 polarization 조합을 처리할 수 있는 네트워크를 고안하였다. 식 (1)은 첫 번째 convolution layer 의 식이다.

$$Y = f(W_{hh} * x_{hh} + W_{hv} * x_{hv} + W_{vh} * x_{vh} + W_{vv} * x_{vv} + b) \quad (1)$$

여기서 $W_{hh}, W_{hv}, W_{vh}, W_{vv}$ 는 3D 필터를 각각 나타낸다. 그리고 *는 3D 콘볼루션 operator 이다. x_{yz} y-z polarization 2D 입력을 나타내고, b 는 bias 항이다. modal 수에 따라 feature map 이 달라질 수 밖에 없는 것을 확인 할 수 있다. 이러한 현상은 네트워크의 성능을 저하시켰다. 따라서 네트워크의 입력을 존재하는 모달의 역수에 비례하는 scale factor 를 도입하였다. 그리고 이 scale factor 를 이용하여 입력의 크기를 조절하였다. 표 2 와 3 은 네트워크가 각각 입력 크기조절을 도입하지 않았을 때와 도입하지 했을 때의 성능을 보여준다.

제안된 입력 크기조절 방법은 멀티모달 데이터에서 큰 성능 개선을 만들었다. 하지만 이러한 성능 향상에도 불구하고, 네트워크의 성능은 특정 polarization 조합만을 처리하는 네트워크의 성능보다 좋지 못하다. 우리는 이러한 현상이 증가된 입력 채널과 다양한 경우의 polarization 존재 여부로 인하여 네트워크가 학습에 어려움을 겪기 때문이라고 가설을 세웠다.

우리는 이러한 학습의 어려움을 완화시키고자 전이 학습 방법을 도입하였다. 전이학습이란 네트워크를 쉬운 일에 먼저 학습 시킨 후 어려운 일에 학습시키는 방법을 의미한다 [6]. 표 4 는 전이학습을 통해 학습시킨 네트워크의 성능을 보여준다. 전이학습을 통하여 네트워크가 학습되었을 때, 네트워크는 특정 모달의 조합 경우만을 위해 학습된 네트워크보다 높은 성능을 보여준다.

그림 2 는 학습된 네트워크들의 filter weight 를 보여준다. 그림 2 에서 확인 할 수 있듯이, 입력 크기조절은 네트워크의 dead filter 의 수를 감소시켰다. 이것은 네트워크가 자신의 capacity 를 충분히 활용한다는 것을 의미한다. 또 제안된 네트워크는 다양한 모달을 활용하여 특징맵을 구성하였고, 이는 네트워크가 모달간의 상관관계를 학습했다는 것을 의미한다.

3. 결론

본 논문은 SAR 영상의 일부 polarization 데이터가 부재 할 수 있는 사실을 바탕으로 SAR 영상을 멀티모달 데이터로 해석하였다. 입력 크기조절을 통하여 모든 polarization 조합을 처리할 수 있는 네트워크를 고안하였다. 제안된 네트워크는 네트워크의 capacity 를 충분히 활용하였고, 멀티모달 학습의 핵심인 모달간의 상관관계도 학습할 수 있었다. 그 결과 입력 크기조절은 멀티모달 데이터에 관하여 큰 성능향상을 만들었다. 제안된 네트워크를 전이학습방법을 통하여 학습시켰을 때, 네트워크는 특정 모달의 조합 경우만을 처리할 수 있는 네트워크보다 더 좋은 성능을 보여주었다.

표 1. 특정 타입의 polarization 조합을 위한 네트워크 성능

Polarization	HH	HV	VH	VV	All
인식률(%)	96.71	97.69	98.90	98.29	99.39

표 2. 입력 크기조절 없는 네트워크 성능

Polarization	HH	HV	VH	VV	All
인식률(%)	92.08	94.15	93.42	92.45	94.03

표 3. 입력 크기조절 네트워크 성능

Polarization	HH	HV	VH	VV	All
인식률(%)	95.74	97.81	98.66	96.59	99.39

표 4. 전이학습을 통해 학습한 입력 크기조절 네트워크 성능

Polarization	HH	HV	VH	VV	All
인식률(%)	97.44	98.42	98.17	98.42	99.51

감사의 글

본 연구는 미래창조과학부 및 정보통신기술진흥센터의 대학 ICT 연구센터육성 지원사업(IITP-2017-2016-0-00288)의 연구결과로 수행되었음.

참고문헌

[1] A. Krizhevsky, S. Ilya S, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097-1105, 2012.

[2] K. Simonyan and A. Zisserman. "Very deep convolutional networks for large-scale image recognition," arXiv: 1409.1556, 2014.

[3] J. Ding, B. Chen, H. Liu and M. Huang, "Convolutional Neural Network with Data Augmentation for SAR Target Recognition," *IEEE Geoscience and Remote Sensing Lett*, Vol. 13, pp. 364-368, Mar, 2016.

[4] S. Chen and H. Wang, "Target Classification Using the Deep Convolutional Networks for SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol.54, pp. 4806 - 4817, Aug, 2016.

[5] L. M. Novak, M. C. Burl, R.D. Chaney and G. J. Owirka, "Optimal Processing of Polarimetric Synthetic-Aperture Radar Imagery", *Lincoln Laboratory Journal*. Vol 3, no.2, pp 273-290, 1990

[6] K. Yu, C. Dong, C. C. Loy, X. Tang, "Deep Convolution Networks for Compression Artifacts Reduction" arXiv: 1608.02778. 2016.