

깊은 합성곱 신경망을 이용한 Synthetic Aperture Radar 영상 내 반전 잡음 성분 제거 기법

김문흠, *이정현, **정제창

한양대학교 전자컴퓨터통신공학과

heum10004@naver.com, *mizali@hanyang.ac.kr, **jjeong@hanyang.ac.kr

A Despeckling Method Using Deep Convolutional Neural Network in Synthetic Aperture Radar Image

Kim, Moonheum *Lee, Junghyun **Jeong, Jaechang

Department of Electronics and Computer Engineering, Hanyang University

요약

본 논문에서는 깊은 합성곱 신경망 (Deep Convolutional Neural Network) 를 이용해서 SAR (Synthetic Aperture Radar) 영상의 반전 잡음 (speckle noise) 성분을 제거하는 기법을 제안하고자 한다. Deep Convolutional Neural Network는 이미지의 데이터 특성에 적합한 딥 러닝 방법이고, 이는 SAR 위성영상의 반전 잡음 제거에 사용해도 효과적이다.

반전 잡음 필터 모델 추정을 위한 학습은 임의로 반전 잡음을 합성한 트레이닝 이미지들과 원본 트레이닝 이미지들을 이용한 회귀모델을 통해 진행된다. 학습을 통해 얻은 반전 잡음 필터는 기존 알고리즘에 비해 우수한 외삽선 보존 성능을 나타냄을 확인하였다.

1. 서론

SAR (Synthetic Aperture Radar) 은 기상조건이나 주야에 관계 없이 영상을 고해상도로 얻을 수 있기 때문에 위성영상취득에 적합한 장비이다. 하지만 라디오파를 지정 위치에 방사한 후 반사된 전파를 합성해 이미지화 하는 SAR장비의 특성상 SAR 이미지는 반전 잡음 (speckle noise) 에 취약하다는 특성이 있다. 반전 잡음은 강한 승산 잡음 (multiplicative noise) 의 일종인데, 이는 위성영상을 이용한 여러 가지 자동화된 알고리즘의 성능을 저해한다. 이에 반전 잡음을 제거하기 위한 다양한 방법들이 제시되었다.[1,2,3,4] 기존에 제시되었던 대부분의 방법들은 반전 잡음의 통계학적인 특성을 이용해서 이미지를 개선하는 방식이었지만, 다양한 분산 값을 가지는 반전 잡음에 대해서는 효과적으로 노이즈를 제거하지 못하는 경향이 있다.

본 논문에서는 딥 러닝 기법 중 하나인 합성곱 신경망 (CNN: Convolutional Neural Network) 을 사용하여 반전 잡음을 제거하고자 한다. 합성곱 신경망은 이미 가우시안 노이즈를 제거하는 데에 높은 성능을 보인다.[5] 노이즈가 가산 잡음(additive noise)일 경우 DnCNN (Denoising Convolutional Neural Network)은 잡음이 섞인 이미지에서 잡음 성분만을 추출하는 방식으로 이미지를 개선한다. 하지만 반전 잡음처럼 승산 잡음인 경우에는 이러한 방법론을 사용할 수 없다. 반전 잡음이 첨가된 이미지에 로그를 취한 후 잡음의 형태를 승산 잡음에서 가산 잡음의 형태로 만들어서 문제를 해결해 보는 것도 고려해볼만하다. 하지만 학습 시 언더 피팅 문제가 생길 경우, 이런 방법은 오히려 로그 역변환시에 원본이미지와 필터링 된 결과

이미지의 오차를 더욱 키울 수 가 있다.

따라서 반전 잡음이 섞인 영상의 경우 합성곱 신경망을 통해서 이미지에서잡음성분만을 추출하는 것보다는 노이즈가 섞인 이미지와 원본이미지를 이용한 회귀모델을 구성하는 것이 반전 잡음 제거 필터를 모델링 하는 데에 효과적이다.

본 논문의 구성은 다음과 같다. 2장에서는 합성곱 신경망을 통해서 잡음 제거 필터를 학습하는 전반적인 이론적 배경을 설명할 것이다. 그리고 3장은 구체적으로 학습을 위한 전처리 과정이나 합성곱 신경망의 중간층의 가중치를 설정하는 방식 등 잡음 제거 필터의 모델링에 대해 구체적으로 설명한다. 4장에서는 제안된 잡음 제거 필터의 성능을 실험하고 분석한다. 마지막으로 5장에서는 결론을 제시한다.

2. 이론적 배경

딥 러닝 기법중의 하나인 합성곱 신경망은 이미지 관련 학습에 특화되어 있는 신경망이다. 합성곱 신경망은 기본적인 인공신경망과는 달리 데이터의 지역적인 부분을 중시하는 것이 특징이다. 이런 방식은 생물이 시각정보를 인식하는 방법과 유사하다. 합성곱 신경망은 특히 잡음 제거에도 좋은 성능을 보이는데 대표적인 것이 가우시안 잡음 제거 신경망인 DnCNN이다.

잡음 제거에 사용되는 합성곱 신경망은 분류나 인식에 사용되는 신경망과는 조금 다른 형태를 가진다. 분류나 인식에 사용되는 합성곱 신경망은 기본적으로 차원 축소를 하는 작업이 존재한다. 차원 축소 과정은 모델의 복잡도를 줄이고 반복되는 정

보 중에서 필요한 정보만을 뽑아내는 역할을 한다. 그리고 분류나 인식을 위해서는 결국 얻어낸 정보들을 이진 비트스트림으로 나타내는 작업이 필요한데 이를 위해서도 차원 축소는 꼭 필요한 작업이다. 하지만 잡음제거 신경망의 경우에는 원하는 결과가 원본 이미지가기에 때문에 이런 차원 축소 과정이 생략 된다. 오히려 컨볼루션을 진행할 때 이미지의 사이즈가 작아지는 것을 방지하기 위해서 필요한 만큼 패딩을 해주는 경우도 있다. 다음은 설명한 두 종류의 CNN의 기본적인 구조이다.

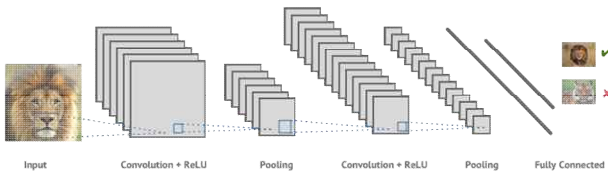


그림 1. 분류, 인식을 위한 CNN 구성[6]

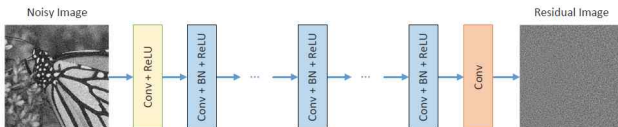


그림 2. 잡음 제거를 위한 CNN 구성[5]

잡음제거는 기본적으로 회귀 모델을 사용한다. 잡음이 섞인 이미지를 모델에 입력했을 때 가장 원본 이미지에 가까운 이미지로 추정해주는 함수를 찾는 것이 잡음 제거에 사용하는 회귀 모델의 핵심이다. 이때 오차를 추정하는 함수의 종류는 다양한데 가장 일반적으로 쓰이는 것은 평균제곱오차이다. 잡음 제거 필터에서 학습의 목표로 사용할 Loss 함수는 평균제곱오차를 사용하는데 그 방법은 식 (1)과 같다.

$$Loss = \frac{1}{N} \sum_{i=1}^N \| f(x_i) - y_i' \|^2 \quad (1)$$

위의 식에서 f 는 CNN을 통해서 추정하는 잡음 제거 필터의 함수이다. 따라서 x_i 는 반전 잡음이 추가된 이미지를 의미한다. y_i' 는 반전 잡음이 섞이기 전의 원본 이미지를 의미하고, 이 Loss function이 줄어든다는 의미는 결국 원본이미지와 CNN 모델을 통과한 이미지의 평균제곱오차가 줄어든다는 말이기 때문에 잡음이 제거된다는 의미와 같다. 마지막으로 N 은 학습에 사용되는 배치의 크기를 뜻한다.

3. 제안하는 방법

3.1 이미지 전처리 과정

입력 데이터 전처리와, 신경망의 가중치나 바이어스 초기화는 신경망 학습의 효율이나 속도를 결정하는 중요한 파라미터이다. 이미지 데이터의 경우에는 보통 한 픽셀 당 8비트의 크기를 가지고 있으므로 0에서 255사이의 값을 가진다. 하지만 이 값들은 보통 신경망의 가중치가 보통 1이하의 실수로 초기화 된다는 것을 고려했을 때 비교적 큰 수이다. 따라서 학습의 효율을 높이기 위해서 이미지 데이터들을 정규화 하는 과정이 필요하다. 제안하는 방법에서는 이 최솟값과 최댓값의

차로 정규화 하는 min-max normalization을 사용하였다. 이론대로라면 min-max normalization을 할 때에는 해당 이미지의 최솟값과 최댓값을 이용하여 정규화를 해야 하지만, 필터 모델의 일관성을 늘리기 위해서, 이미지의 최솟값은 0, 그리고 최댓값은 255로 가정하고 정규화를 한다. 때문에 정규화 작업의 결과는 모든 픽셀을 255로 나누는 것과 동일하다고 볼 수 있다. 정규화 과정은 학습에 사용되는 데이터 전부에 동일하게 적용된다. 신경망 모델을 통과하지 않는 원본 이미지에도 적용하는 이유는 학습 진행시 평균제곱오차의 크기를 낮춰서 학습의 효율을 증가시키기 위함이다.

3.2 학습 데이터 구성

학습에 사용되는 데이터는 여러 가지 이미지에 반전 잡음을 섞어서 만들었다. 다만 이미지 원 크기 그대로 학습을 진행하면 그에 비례해서 신경망의 가중치 개수가 증가하고, 모델의 복잡도가 너무 높아지기 때문에 반전 잡음이 추가된 여러 이미지에서 무작위로 64x64픽셀 크기의 패치들을 추출해서 학습데이터를 구성했다.

그리고 원본 데이터에서도 동일한 위치에서 패치를 추출해서 목표 이미지 데이터 셋을 구성하였다.



그림 3. 데이터 셋 예시

모델을 학습시키는 데에 사용한 이미지 패치는 학습용 데이터 셋, 그리고 목표 데이터 셋 각각 1,000장을 사용하였고, 밸리데이터 셋 (validation set) 과 테스트 셋에는 각각 400장 100장을 사용하였다.

학습 진행시의 배치 구성은 모델의 오버 피팅을 예방하기 위해서 에포크(epoch) 당 배치사이즈로 20개의 이미지 패치를 사용하였고, 20장의 이미지는 1,000개의 학습용 데이터 패치들 중 에서 무작위로 추출된 패치를 사용하였다.

3.3 신경망 구성

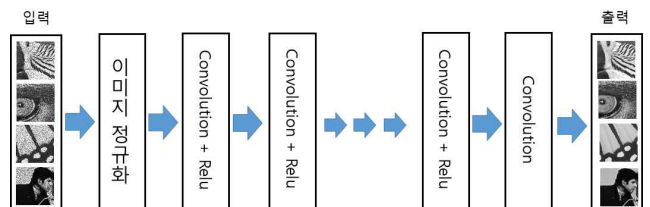


그림 4. 제안한 신경망의 구성

이미지 패치들은 정규화 과정을 거치고 합성곱 신경망에 입력된다. Deep CNN의 경우 Relu(Rectified linear unit)가 여타 다른 함수의 활성화 함수들 보다 좋은 성능을 보이기 때문에, 제안한 필터 모델에도 활성화 함수로 Relu를 사용하였다. 컨볼루션 계층은 총 7개로 구성하였는데, 각 층의 커널의 개수는 16, 32, 64로 층이 두 개 증가할 때마다 커널의 개수도 두 배로 증가하는 형식으로 구성하였다. Loss 함수로는 앞에서 설명한 바와 같이 제곱오차함수를 이용하였고, optimizer로는 adam-optimizer (adaptive moment estimation optimizer)를 사용하였다. Optimizer에 사용된 학습률 (learning rate)은 1×10^{-5} 이다.

3.4 이미지 복원

위의 학습된 CNN은 64x64픽셀 크기로 학습했기 때문에 64x64픽셀 크기의 이미지에만 적용된다. 따라서 패치크기보다 큰 이미지를 복원할 때는 이미지를 스캔하듯이 필터를 옮겨가면서 이미지를 모델에 통과시키면 복원된 이미지를 얻을 수 있다. 이때 결과이미지는 min-max normalization된 결과이기 때문에 결과이미지의 각 픽셀에 255를 곱해 주어야 한다.

4. 실험 결과 및 분석

본 실험은 Window embedded 8 64 비트 운영체제, CPU: Intel(R) Core(TM) i7-4790CPU 3.60Hz, GPU: GTX 960 의 PC환경에서 python, tensorflow를 이용하여 실험을 진행하였다. 학습은 제안한 신경망과 학습데이터를 이용해 50,000회 진행하였다. 다음은 잡음 제거 필터를 학습한 신경망을 이용해서 두 개의 위성사진에서 반전 잡음을 제거한 결과이다.



그림 5. 반전 잡음이 섞인 노이즈 이미지



그림 6. 반전 잡음을 제거한 이미지



그림 7. 원본 이미지

그림 6 과 그림 7의 2개 이미지를 통해 주관적 화질을 비교해보면, 제안한 반전 잡음 필터가 반전 잡음을 제거함을 확인할 수 있었다. 하지만 외곽선 성분이 약한 디테일은 제대로 복원하지 못하고 블러링(blurring) 현상이 나타나는 것 또한 확인할 수 있었다. 또한 바다와 강과 같이 비슷한 픽셀 값이 연속되는 구간에는 필터의 복원이 일정하지 못하여 구획화 현상(blocking artifact)이 발생하는 것을 확인되었다.

다음은 실험 결과 이미지와 원본이미지의 객관적인 화질을 PSNR (Peak Signal to Noise Ratio)과 SSIM(Structural Similarity)을 통해 기존의 잡음 제거 필터들과 비교해 보았다.

	noised image	bilateral filter	guided filter	proposed
PSNR (dB)	21.3272	23.3692	24.1517	27.2118
SSIM	0.6507	0.7323	0.7535	0.8297

표 1. 첫 번째 위성사진의 객관적 화질 실험 결과

	noised image	bilateral filter	guided filter	proposed
PSNR (dB)	20.9841	21.5207	22.0265	22.8310
SSIM	0.8399	0.8527	0.8613	0.8834

표 2. 두 번째 위성사진의 객관적 화질 실험 결과

표 1과 표2를 통해 제안된 반전 잡음 필터가 PSNR과 SSIM측면에서 기존의 잡음 제거 필터들보다 개선된 성능을 보인다는 것을 확인할 수 있었다. 기존의 잡음 제거 필터중 가장 좋은 성능을 보이는 guided filter와 비교하여 첫 번째 사진의 경우 PSNR이 3.0601dB 개선되었고, 두 번째 사진의 경우 PSNR이 0.8045dB만큼 개선되었다. 첫 번째 사진의 경우 평평한 영역의 분포가 많이 포함되면서, 잡음 제거의 효과가 PSNR측면에서 특히 개선되었다.

SSIM 수치 측면에서 잡음 제거 필터와 제안하는 방법의 결과를 비교했을 때, 첫 번째 위성사진의 경우 guided filter와 비교하여 0.0762만큼 더 높은 수치를 보였고, 두 번째 위성사진의 경우 0.0221만큼 더 높은 수치를 보였다. PSNR 측면의 비교와 마찬가지로 첫 번째 위성사진의 경우 기존의 필터들보다 특히 개선된 SSIM수치를 보였다. 두 번째 위성사진처럼 외곽선 성분이 강한 이미지가 아님에도 SSIM 수치가 높게 나오는 것을 근거로 제안된 반전 잡음 필터가 기존의 잡음 제거 필터보다 우수한 외곽선 보존 성능을 보인다는 것을 알 수 있다.

5. 결론

본 논문에서는 합성 곱 신경망 학습을 통한 반전 잡음 제거 필터를 제안 하였고, 제안된 필터는 기존의 알고리즘들과 비교했을 때 개선된 결과를 보였다. 특히 실험결과에 따르면 합성 곱 신경망 학습을 통한 반전 잡음 제거 필터는 외곽선 보존 성능이 우수한 것으로 나타났다. 하지만 주관적 화질로 비교해볼 때 블러링, 구획화 현상 등의 단점이 있었다. 이 단점의 개선을 위해 신경망 학습을 최적화 하고, 이미지의 디테일을 보존할 수 있는 전처리 과정이 필요함을 알 수 있다.

감사의 글

“본 연구는 미래창조과학부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음.[2014-0-00670, ICT 장비용 SW 플랫폼 구축]”

참고문헌

- [1] F. Argenti, A. Lapini, L. Alparone, and T. Bianchi, “A tutorial on speckle reduction in synthetic aperture radar images,” *IEEE Geosci. Remote Sens. Mag.*, vol. 1, pp. 6 - 35, 2013.
- [2] T. Bianchi, F. Argenti, and L. Alparone, “Segmentation-Based MAP Despeckling of SAR Images in the Undecimated Wavelet Domain,” *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2728 - 2742, 2008.
- [3] S. Foucher, “SAR image filtering via learned dictionaries and sparse representations,” in *IEEE IGARSS*, 2008, pp. 229 - 232.
- [4] C.-A. Deledalle, L. Denis, and F. Tupin, “Iterative weighted maximum likelihood denoising with probabilistic patch-based weights,” *IEEE Trans. on Image Process.*, vol. 18, no. 12, pp. 2661 - 2672, 2009.
- [5] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising,” *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142 - 3155, Jul. 2017.
- [6] <https://shafeentejani.github.io/2016-12-20/convolutional-neural-nets/>