

# 의사 형태학적 연산을 사용한 이미지 변환

조장훈, 이호연, 신명우, 김경섭  
충남대학교 컴퓨터공학과

e-mail : everestbeta524@gmail.com , kiwi3618@naver.com , qew9191@naver.com ,  
sclkim@cnu.ac.kr

## Image Translation using Pseudo-Morphological Operator

Janghun Jo, HoYeon Lee, MyeongWoo Shin, Kyungsup Kim  
Dept. of Computer Engineering, Chung-Nam National University

### 요 약

이 연구에서는 형태학적 연산(Morphological Operator)과 CNN (Convolutional Neural Networks)의 개념을 결합하여 이미지 변환을 개선하고자 한다. 이를 위해서 형태학적 연산을 근사할 수 있는 연산을 제안한다. 그리고 제안한 연산을 CNN처럼 여러 필터를 사용할 수 있게 확장한 S-Convolution을 제안한다. 실험 결과 제안한 연산은 형태학적 연산을 학습할 수 있었다. 그리고 제안한 연산의 이미지 변환 성능을 검증하기 위해 GAN에 적용하여 실험하였다. 그 결과 S-Convolution이 기존 CNN을 사용한 GAN과 다른 변환이 가능하다는 것을 볼 수 있었다.

### Abstract

We attempt to combines concepts of Morphological Operator(MO) and Convolutional Neural Networks(CNN) to improve image-to-image translation. To do this, we propose an operation that approximates morphological operations. Also we propose S-Convolution, an operation that extends the operation to use multiple filters like CNN. The experiment result shows that it can learn MO with big filter using multiple S-convolution layer of small filter. To validate effectiveness of the proposed layer in image-to-image translation we experiment with GAN with S-convolution applied. The result showed that GAN with S-convolution can achieve distinct result from that of GAN with CNN.

### 1. 서론

영상처리의 여러 문제는 이미지 변환 문제로 정의할 수 있다. 최근 이미지 변환 문제를 딥러닝을 이용해서 해결하려는 여러 시도가 있었다. 많은 연구자가 딥러닝을 활용해서 이미지를 변환하는 신경망 모델을 학습시키는 방법을 사용하는데, 주로 CNN (Convolutional Neural Networks), Transposed Convolution (Deconvolution), VAE(Variational Autoencoder), GAN (Generative Adversarial Networks) 등이 사용된다[1]. 그 중 GAN 을 사용한 접근법이 이미지 변환에서 좋은 성능을 보였다[2][7].

형태학적 영상처리는 수학적 형태학(Mathematical Morphology)에 기반을 둔 분야이다. 형태학적 영상처리에는 구조요소(Structuring element)라는 필터를 사용하는 팽창(Dilation), 침식(Erosion)의 두 연산이 있으며 각각 최댓값 최솟값 연산에 대응한다. 이 연산은 비선형이며 노이즈 제거, 골격화, 물체 검출 등에 사용된다[9]. 수학적 형태학의 비선형필터는 CNN 으로 학습하기가 어렵다. 다른 연구에서는 형태학적 연산자를 학습시키기 위해 형태학적 연산을 근사할 수 있는 반 조화평균(Contra-Harmonic Mean)을 이용한 반

조화평균 필터(Contra-Harmonic Mean Filter)를 사용하였다[4][11].

본 논문에서는 GAN 에 형태학적 연산을 적용해서 이미지 변환을 시도한다. 기존 GAN 은 CNN 을 사용했기 때문에 형태학적 연산자를 사용하면 기존과는 다른 이미지 변환 결과를 얻을 것으로 기대된다. 형태학적 연산자를 그대로 사용하면 연산 자체를 학습할 수 없기 때문에 형태학적 연산 자체를 학습할 수 있는 함수를 사용한다. 그리고 반 조화평균 필터는 입력이 양수에서만 정의되기 때문에 이 제한을 피하고자 반조화평균이 아닌 Smooth Maximum[5]이라는 함수를 사용해 유사 형태학 연산자를 정의한다. 이 연산을 사용하면 형태학적 연산을 학습할 수 있을 뿐만 아니라 max, min 풀링(Pooling), 제한적인 컨볼루션을 표현할 수 있다. 그리고 이 연산을 CNN 처럼 여러 필터를 사용할 수 있게 확장한 S-Convolution 을 정의한다.

이 함수를 딥러닝 모델에 적용하여 형태학적 연산 및 일반적인 이미지 변환의 학습이 가능함을 검증하기 위해. 제안한 연산을 사용하여 형태학적 연산의 학습을 시도한다. 그리고 기존 딥러닝 모델의 생성망

<sup>1</sup> 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW 중심대학지원사업의 연구결과로 수행되었음 (2015-0-00930)

(Generator) 부분에 사용하여 이미지 변환을 시도했다. 그 결과는 기존 딥 러닝 모델보다 물체의 경계 부분을 더 명확히 학습하는 것으로 나타났다.

## 2. 연구 배경

### 2.1 Generative Adversarial Networks

GAN은 게임 이론을 인공 신경망에 적용해서 실제 데이터의 분포를 학습하는 신경망이다. 생성망 (Generator)과 구별망(Discriminator)의 두 인공신경망으로 구성된다. 생성망은 가짜 데이터를 생성하며 구별망은 생성된 가짜 데이터와 진짜 데이터를 구별한다. 여기서 생성망은 가짜 데이터가 구별망에서 진짜로 분류되도록 학습을 시키고, 구별망은 진짜, 가짜 데이터를 더 잘 분류하게 학습을 시킨다. 이는 생성망과 구별망의 minimax 게임으로 표현될 수 있다. 학습이 진행되면 생성망이 생성한 가짜 데이터는 진짜 데이터의 분포와 유사해진다[1].

Conditional GAN(cGAN)은 GAN이 생성하는 데이터를 조절할 수 있게 변형한 신경망이다[6]. GAN에 추가로 라벨[6], 텍스트[3], 이미지[2][7] 등을 주어 생성물을 조절한다. 이를 이용해서 이미지를 변환하는 방법이 제안되었다[2][7].

### 2.2 Smooth Maximum

Smooth maximum 함수는 최댓값(Max), 최솟값(Min) 함수를 미분가능하게 근사하는데 사용되는 함수이며 다음과 같이 정의된다[5].

$$S_{\alpha}(x) = \frac{\sum_{i=1}^n x_i e^{\alpha x_i}}{\sum_{i=1}^n e^{\alpha x_i}} \quad (\text{식. 1})$$

이 함수는 다음과 같은 성질을 갖는다.

- 1)  $\lim_{n \rightarrow \infty} S_{\alpha}(x) = \max(x)$
- 2)  $S_0(x) = \frac{\sum_{i=1}^n x_i e^{0 \cdot x_i}}{\sum_{i=1}^n e^{0 \cdot x_i}} = \frac{1}{n} \sum_{i=1}^n x_i$
- 3)  $\lim_{n \rightarrow -\infty} S_{\alpha}(x) = \min(x)$

### 3. Smooth maximum Convolution

Weighted-Smooth maximum은 Smooth maximum에 가중치  $w$ 를 추가한 함수를 다음과 같이 정의한다.

$$S_{\alpha, w}(x) = \frac{\sum_{i=1}^n w_i x_i e^{\alpha x_i}}{\sum_{i=1}^n w_i e^{\alpha x_i}} \quad (w_i \geq 0) \quad (\text{식. 2})$$

Weighted-Smooth maximum을 이미지 필터로 사용하면  $\alpha = 0$ 일 때에는 필터가  $\frac{w_i}{\sum_{i=1}^n w_i}$ 인 컨볼루션이 된다. 그리고  $\alpha > 0$  이면 확장 연산,  $\alpha < 0$  이면 침식 연산을 근사할 수 있다. 이 때 구조요소는  $\ln w_i$ 로 주어진다. 그럴 뿐만 아니라 이 함수는 max, min 함수를 근사할 수 있기 때문에 풀링(Pooling) 연산을 할 수도 있다.

CNN은 여러 개의 필터를 사용해 채널간 연산을 할 수 있다. 비슷하게 Weighted-Smooth maximum을 확장한 연산 및 계층을 S-Convolution이라고 정의한다. S-Conv(S-Convolution)은 기존 컨볼루션과 같이 여러 계층을 구성하여 딥 네트워크에 적용할 수 있을 뿐만 아니라, 여러 간의 연산이 가능하다.

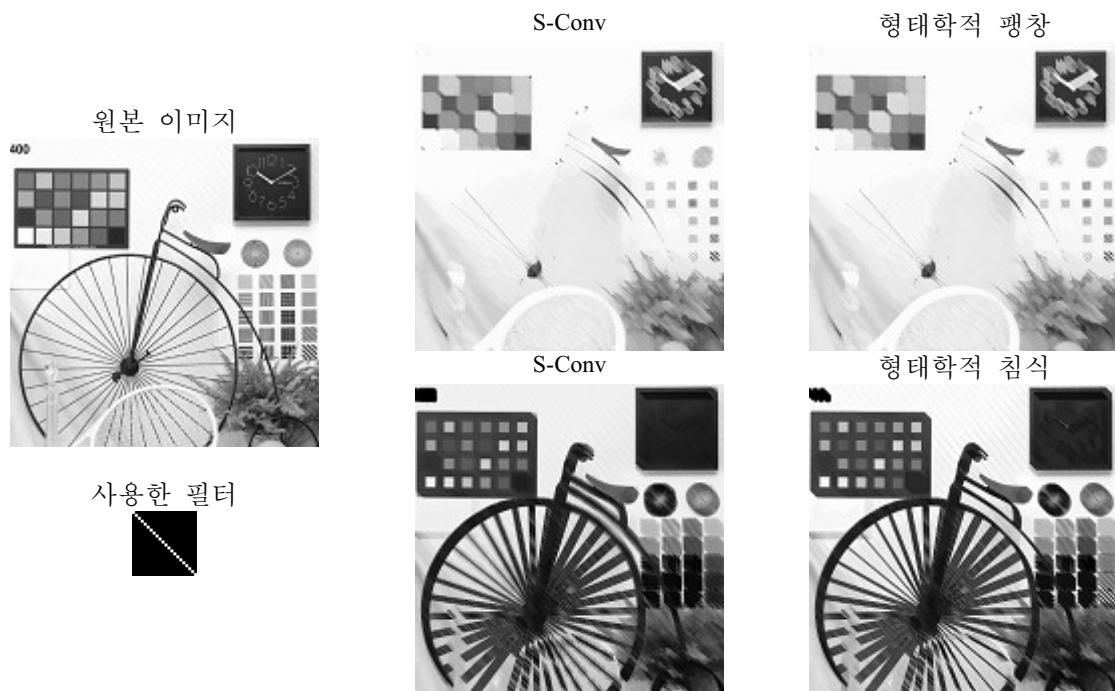


그림1. S-Conv로 형태학적 연산을 학습한 결과

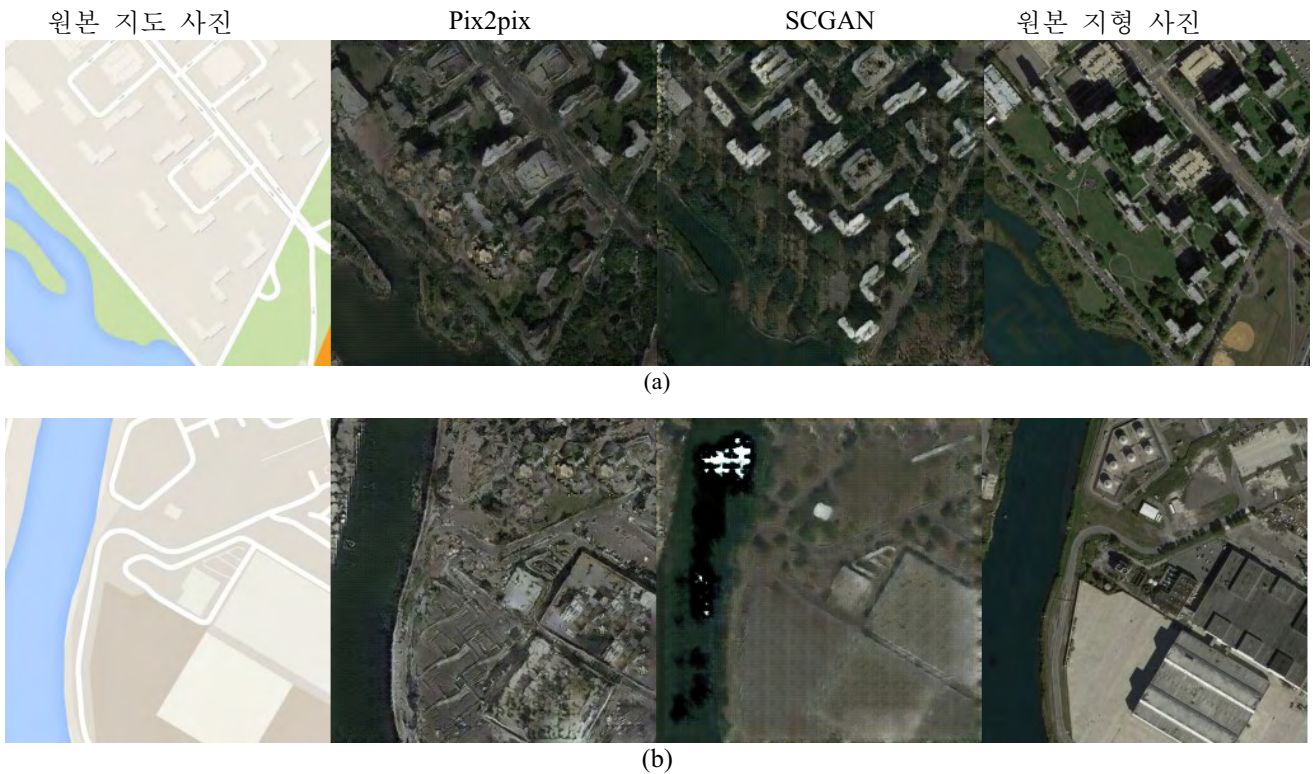


그림3. 지도 사진을 입력으로 지형 이미지를 생성한 예

한편 이미지 변환에는 전치컨볼루션(Transposed Convolution)[8]이라는 연산이 사용되는데 이는 컨볼루션 연산의 역연산 형태로써 컨볼루션이 입력에 필터를 곱하여 결과를 생성한다면, 전치컨볼루션은 결과를 필터로 나누어 입력을 구하는 연산이다. 이미지 변환에 사용하기 위해 S-Conv를 전치 S-Conv(Transposed S-Convolution)로 확장한다. 이는 S-Conv 연산에서 컨볼루션 부분을 전치컨볼루션으로 대체한 연산으로 정의한다.

S-Conv가 형태학적 연산자를 학습할 수 있는 것을 검증하기 위해 11x11의 구조요소의 형태학 연산의 학습을 시도했다. 원본 이미지[10]에 대각선 필터를 사용해 확장, 침식을 적용한 이미지를 생성해서 학습에 사용했다. 손실 함수는 평균제곱오차를, 학습 속도는 0.001에 1000번 반복마다 0.96의 감쇠율을 적용하고 학습 횟수는 10000번, 배치 크기는 1으로 설정하고 최적화는 Adam을 사용하였다.

11x11의 필터를 가진 단일 S-Conv 계층으로 학습한 결과 형태학적 연산을 잘 학습하였다. 그리고 계층이 많은 경우의 성능을 확인하기 위해 5x5의 필터를 가진 S-Conv 계층 3개로 이루어진 네트워크를 사용해 학습을 시도했다. 그 결과는 그림 1에서 확인할 수 있는데 확장 연산은 학습이 매우 잘 되었지만 침식 연산은 학습이 비교적 어려웠다.

#### 4. SCGAN (S-Convolution GAN)

이미지 변환에 S-Conv를 적용하기 위해 기존 GAN인 Pix2Pix[2]에 S-Conv를 적용한다. 편의상 S-Conv를 적용한 네트워크를 SCGAN으로 정의한다. Pix2Pix에서 변경한 부분은 생성망의 디코더(Decoder) 부분으로, 디코더의 마지막 계층은 그대로 두고 디코더의 나머지 계층, 즉, 9~15번째 계층을 전치-S-Conv로 대체했다. 그림 2는 네트워크 구조를 간략히 한 것이다. 훈련에 사용하는 변수들은 Pix2Pix와 동일하게 설정하였다.

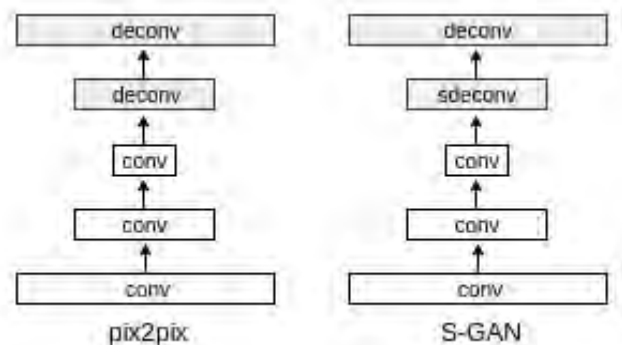


그림 2. Pix2pix와 SCGAN의 생성망 계층

학습 과정에서는 학습 이미지로 Pix2Pix의 map2sat를 사용하여 지도 사진을 위성 사진으로 변환시키도록 하였다. 그 결과는 그림 3에서 확인할 수 있다. 그림 3.(a)에서 일반적인 건물과 도로가 있는 지도를 사용하여 변환할 경우에는 좀 더 선명하게 변환하는 것으로 보이지만 그림 3.(b)에서는 물이나 숲이 많은 경우에는 변환하는데 어려움이 있는 것으로 보인다.

Pix2Pix와 비교하면 Pix2Pix는 지도의 도로나 건물 등의 윤곽선에 잘 맞추지 못하고 벗어나게 변환하는 것에 비하여, SCGAN은 윤곽선을 잘 맞추어서 변환해내는 것을 볼 수 있다. 하지만 지나치게 윤곽선을 중심으로 하여 이미지를 생성하는 모습을 보였는데, 필터 크기 조절이나 구조 변경을 통해 수정하면 이러한 문제도 해결될 것으로 예상된다.

## 5. 결론

본 논문에서는 Smooth Maximum을 확장하여 형태학적 연산을 근사하는 함수를 제안하였다. 그리고 이를 딥러닝에 적용한 S-Conv를 제안했다. 실험 결과 S-Conv는 단일 또는 여러 계층으로 형태학적 연산을 학습할 수 있었다. 기존 GAN을 이용한 이미지 변환에 S-Conv를 사용한 결과 컨볼루션을 사용한 기존의 GAN에 비해 출력 이미지의 구조가 원본 이미지의 구조와 유사하게 개선되었다. 이는 S-Conv가 형태학적 연산의 특성을 가져서 구조적인 사물을 더 잘 표현하기 때문으로 보인다. 향후 S-Conv를 다른 네트워크에도 적용할 수 있을 것으로 보인다.

## 참고문헌

- [1] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," Jun. 2014.
- [2] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," Nov. 2016.
- [3] Zhang, Han, et al. "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks." *arXiv preprint arXiv:1612.03242*(2016).
- [4] J. Masci, J. Angulo, and J. Schmidhuber, "A Learning Framework for Morphological Operators using Counter-Harmonic Mean."
- [5] M. Lange, D. Zühlke, O. Holz, and T. Villmann, "Applications of  $l_p$ -Norms and their Smooth Approximations for Gradient Based Learning Vector Quantization."
- [6] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," Nov. 2014.
- [7] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *arXiv preprint arXiv:1703.10593* (2017).
- [8] Zeiler, Matthew D., et al. "Deconvolutional networks." *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010.
- [9] McAndrew, Alasdair. "An introduction to digital image processing with matlab notes for scm2511 image

processing." *School of Computer Science and Mathematics, Victoria University of Technology* (2004): 1-264.

- [10] "dataset of standard 512x512 grayscale test images." [Online]. Available: <http://decsai.ugr.es/cvg/CG/base.htm>. [Accessed: 17-Sep-2017].
- [11] J. Angulo, "Generalised morphological image diffusion," *Nonlinear Anal. Theory, Methods Appl.*, vol. 134, pp. 1–30, 2016.