

폴립 영역 분할을 위한 DIP 기반의 이미지 초해상화 및 데이터 증강 기법에 관한 연구

김현진^o 최중환 이지은 박상현[†]

연세대학교 컴퓨터과학과

{hjkim99, mathcombio, jieun199624, sanghyun}@yonsei.ac.kr

A Study of DIP-based Image Super-resolution and Data Augmentation Methods for Polyp Segmentation

Hyunjin Kim^o Jonghwan Choi Jieun Lee SangHyun Park[†]

Dept. of Computer Science, Yonsei University

요 약

의료 영상 이미지 영역 분할을 위한 초해상화 기술은 병변의 세부적인 구조를 보다 정확하게 분석할 수 있도록 이미지 해상도를 개선하는 이미지 전처리 기법 중 하나이다. 의료 영상 이미지 초해상화를 위한 심층 신경망 모델은 여럿 제안되었으나, 이들을 훈련하는데 필요한 고해상도의 의료 영상 이미지를 수집하는 것이 쉽지 않기 때문에 실용성을 고려한 새로운 이미지 전처리 기술 개발이 필요한 실정이다. 본 연구에서는 폴립 영상 분할 모델의 성능을 향상시키기 위한 이미지 초해상화 및 데이터 증강 기법을 제안한다. 제안 기법은 deep image prior 기술을 활용하여 다량의 훈련데이터 없이 단일 저해상도 이미지만으로 초해상화 및 데이터 증강을 수행한다. 제안하는 방법은 대장 내시경 이미지에 대해 기존 보간법들보다 우수한 초해상화 성능을 보여주었으며, 나아가 폴립 영역 분할 모델의 정확도를 향상시키는 것을 확인하였다.

1. 서 론

단일 이미지 초해상화(single image super resolution; SISR)는 저해상도 이미지를 고해상도 이미지로 복원하기 위한 이미지 처리기술을 가리킨다[1]. 이미지 초해상화는 의료 내시경 영상 분석, 인공위성 영상 분석, 스마트시티 보안카메라 기술 개발 등 광범위한 이미지 분석 기술에서 사용된다[2].

전통적인 초해상화 기술은 저해상도 이미지의 화소 정보를 이용하여 새로운 픽셀(pixel)의 화소값을 추정하는 보간법을 활용한다[3]. 보간법 기반의 초해상화 기술에는 주변의 저해상도 화소값으로 보간을 수행하는 최근접 화소 보간법(nearest neighbor interpolation)과 주변 4개의 화소 값을 통해 3차원 곡선 상의 값으로 보간을 수행하는 양입방 보간법(bicubic interpolation)이 있다. 이러한 보간법 기반의 초해상화는 계산량이 적다는 장점이 있지만 이미지의 세부 영역에 대한 화질 개선 성능이 낮은 한계점이 있다.

합성곱 신경망(convolutional neural network; CNN)과 같은 심층학습 기술의 발달은 고화질의 초해상화 기법 개발을 촉매하였으며, 기존의 보간법에 기반한 초해상화의 한계점을 극복할 수 있었다.

많은 심층학습 모델들이 초해상화 기능을 수행하기 위해 저해상도와 고해상도 이미지 간의 통계적 관계를 분석 및 학습한다. Dong et al.[4]은 3개의 합성곱 모듈, 패치 추출(patch extraction), 비선형 변환(non-linear mapping), 복원(reconstruction)을 구현 및 통합하여 양입방 보간법보다 뛰어난 성능의 신경망 모델 SRCNN을 제안하였다. 단순 CNN 기반의 모델을 개선하기 위해, Lim et al.[6]은 잔차 신경망(residual neural network)으로 이루어진 향상된 심층 초해상화(enhanced deep super resolution; EDSR) 모델을 제안했으며, Zhang et al.[6]은 채널 주의집중 메커니즘을 적용한 잔차 채널 주의집중 신경망(residual channel attention network; RCAN)을 제안하는 등 다량의 학습데이터에 기반한 심층 신경망 모델들이 여럿 제안되었다.

하지만, 막대한 학습데이터가 요구되는 초해상화 기술은 학습데이터 수집이 어려운 의료 영상 분야에서 그 성능을 발휘하는데 한계가 있다. 일반적으로 의료 영상의 공간 해상도는 영상 획득 시간, 낮은 조사량 또는 하드웨어 한계와 같은 제약으로 인해 충분하지 않다[1]. 고가의 영상 획득 장치를 사용하여 고해상도 이미지를 얻을 수 있으나, 그러한 데이터를 수집하는데 고비용이 요구되고, 신호 대 잡음비(signal to noise ratio; SNR)가 낮아진다.

최근의 연구에서는 별도의 학습데이터 없이 곧바로 초해상화 기능을 수행할 수 있는 신경망 모델 훈련 전략인 deep image prior (DIP)가 제안되었다[7]. DIP는 주어진 단일 이미지만을 가지고서 해당 이미지의 고해상도 버전을 확률적으로 생성하기 위해 자기지도 학습(self-supervised learning)을 수행한다. DIP 저자들은 이러한 학습 전략이 초해상화 이외에도 inpainting

* 이 논문은 2023년도 정부(과학기술정보통신부)의 재원으로 정보통신기획지원의 지원(IITP-2017-0-00477, (SW 스타랩) IoT 환경을 위한 고성능 플래시 메모리 스토리지 기반 인메모리 분산 DBMS 연구개발)과 국토교통부의 스마트시티 혁신인재육성사업으로 지원을 받아 수행된 연구임.

† 교신 저자 sanghyun@yonsei.ac.kr

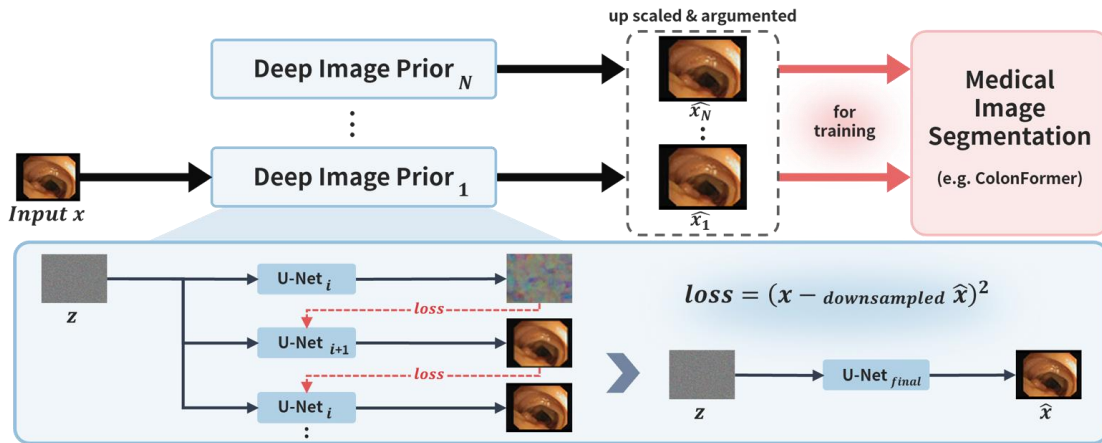


그림 1. MEDIPAM 구조 및 폴립 영상 분해 파이프라인

작업에서도 활용될 수 있음을 실험결과를 통해 보여주었다. 해당 실험결과를 통해 DIP는 학습데이터를 확보하기 어려운 의료 영상 분석에서 효과적인 도구로 자리 매김할 것을 기대할 수 있으며, 특히 초해상화 기술로써 DIP 활용은 내시경 이미지 영역 분할과 같은 작업의 성능을 향상시키는데 크게 기여할 수 있다.

본 연구에서는 대장 내시경 이미지에 대한 폴립 영역 분할(polyp segmentation)을 위해 개발된 심층 신경망 모델의 성능을 효과적으로 향상시키기 위한 DIP 기반의 이미지 초해상화 및 데이터 증강 기법 MEDIPAM (MEDical Deep Image Prior Augmentation)을 제안한다. 제안하는 방법은 단일 저해상도 이미지에 대하여 반복적인 DIP 적용을 통해 복수 개의 고해상도 이미지들을 생성하여 초해상화 및 학습데이터 증강을 동시에 달성한다.

2. 모델 구조

본 연구에서는 향상된 폴립 영역 분할 작업을 위해 U-Net[8] 구조를 갖는 MEDIPAM 및 최신의 폴립 영역 분할 모델인 ColonFormer[9]를 결합하여 사용하였다(그림 1). 다음의 내용은 각각의 모델에 대한 설명이다.

2.1. MEDIPAM

제안하는 모델은 단일 저해상도 이미지에 대하여 DIP를 반복적으로 적용하여 무작위로 생성된 복수개의 고해상도의 이미지를 생성하는 skip-connection을 가진 U-Net 구조[8]의 신경망 모델이다. DIP는 새로운 고해상도 이미지 한 장을 생성할 때마다 초기화된 신경망 모델을 요구하며, 해당 신경망 모델은 균일하게 분포된 무작위 노이즈를 입력 받아서 고해상도 이미지를 출력한다. 출력된 고해상도 이미지로부터 신경망 모델 훈련을 위한 손실함수를 계산하기 위해, 출력 이미지를 다운스케일링(down-scaling)하고 이를 입력 저해상도 이미지와의 비교하여 손실함수를 계산 및 역전파(backpropagation) 알고리즘을 수행한다. DIP는 하나의 신경망 모델을 통해 하나의 고해상도 이미지를 생성할 수 있으며, 신경망 모델의 초기값에

따라 생성된 결과가 달라지기 때문에 단일 저해상도 이미지에 N 번의 DIP 적용은 N 개의 다양한 고해상도 이미지를 수집할 수 있게 한다. MEDIPAM에서 DIP에 대한 손실함수는 두 이미지의 각 픽셀들 간의 L2거리 합으로 계산된다.

2.2. Polyp Segmentation Model

본 연구에서는 대장 내시경 영상 이미지에서 폴립 영역을 효과적으로 분할하기 위해 ColonFormer[9]를 사용하였다. ColonFormer는 장거리 시맨틱 정보를 모델링할 수 있는 인코더-디코더 구조이다. 인코더는 다중 스케일에서 전역 의미론적 관계를 모델링하기 위한 트랜스포머 구조를 가지고 있으며, 디코더는 특징 표현을 풍부하게 하기 위해 다단계 표현을 학습하도록 설계된 계층적 네트워크 구조로 설계되었다.

3. 실험 및 결과

3.1. 벤치마크 데이터집합

본 연구에서는 MEDIPAM의 성능을 평가하기 위해 2개의 대장 내시경 데이터집합 Kvasir와 CVC-ClinicDB를 사용하였다. Kvasir는 노르웨이 Vestre Viken Health Trust (VV)에서 내시경 장비를 사용하여 수집되었으며, VV와 소화기과 전문의가 이미지에 주석을 달고 검증한 1,000개의 이미지로 구성되어 있다. CVC-ClinicDB는 대장 내시경 비디오에서 추출한 이미지 프레임으로 31개의 대장 내시경 시퀀스에서 612개 이미지로 구성되어 있다. 본 연구에서는 Dut et al.[9]에서 제안한 것과 동일한 실험 데이터셋 분할을 사용하며, Kvasir 및 ClinicDB의 90%는 훈련데이터로, 나머지 10%는 시험데이터로 사용되었다.

3.2 실험 결과

MEDIPAM의 성능 평가를 위해 두 가지의 전통적 초해상화 기법인 최근접 화소 보간법과 양입방 보간법을 베이스라인 모델로 사용하였다. 그리고 MEDIPAM의 데이터 증강 횟수에

표 1. 업스케일 팩터에 따른 이미지 크기

| Up-scaling factor | Image size |
|-------------------|------------|
| original | 48 x 36 |
| x2 | 96 x 64 |
| x4 | 192 x 128 |

표 2. 초해상화 의료 영상 분할 실험 결과

| Model | | Dataset | | | |
|-------------|-----|---------|-------|--------------|-------|
| | | Kvasir | | CVC-ClinicDB | |
| Method | USF | Dice | IOU | Dice | IOU |
| Nearest-up | x2 | 0.877 | 0.800 | 0.878 | 0.798 |
| Bicubic-up | | 0.873 | 0.792 | 0.859 | 0.781 |
| MEDIPAM(x1) | | 0.899 | 0.838 | 0.911 | 0.857 |
| MEDIPAM(x2) | | 0.902 | 0.842 | 0.920 | 0.869 |
| MEDIPAM(x3) | | 0.903 | 0.845 | 0.923 | 0.873 |
| Nearest-up | x4 | 0.872 | 0.794 | 0.868 | 0.787 |
| Bicubic-up | | 0.873 | 0.794 | 0.865 | 0.787 |
| MEDIPAM(x1) | | 0.899 | 0.839 | 0.906 | 0.851 |
| MEDIPAM(x2) | | 0.902 | 0.843 | 0.921 | 0.870 |
| MEDIPAM(x3) | | 0.903 | 0.846 | 0.928 | 0.878 |

따른 성능 또한 같이 비교 분석되었다. 성능 평가 절차는 다음과 같다. 훈련데이터에 대해서는 각 초해상화 기법을 적용하여 고해상도로 구성된 훈련데이터를 구축하고, 해당 고해상도 이미지들을 이용하여 ColonFormer를 학습시켰다. 각 초해상화 기법이 ColonFormer의 학습에 얼마나 긍정적인 영향을 주었는지를 알아보기 위해, 시험데이터는 원본 그대로 입력되어 그 정확도를 측정하였다. 초해상화에 관한 업스케일 팩터(up-scaling factor; USF)는 2와 4만 적용되었으며, 표 1은 원본 이미지의 크기 및 업스케일된 고해상도 이미지들의 크기를 보여준다. 폴립 영역 분할 성능은 두 가지 정량적 지표 Dice와 IOU를 통해 측정되었다.

표 2는 두 벤치마크에서 각 초해상화 기법들이 적용된 ColonFormer의 정확도를 보여준다. MEDIPAM(x1)은 데이터 증강없이 초해상화만 적용한 경우를 나타내며, 최근접 화소 보간법(nearset-up) 및 양입방 보간법(bicubic-up) 대비 USF=2일 때 0.033%, USF=4일 때 0.046% 향상된 성능을 보여주었다. 그림 2는 각 초해상화 기법으로 얻은 고해상도 이미지 예시를 보여주며, DIP를 활용한 MEDIPAM이 폴립 부분을 보다 선명하게 보여주는 것을 확인할 수 있다.

MEDIPAM은 초해상화 뿐만 아니라 데이터 증강 또한 가능케 하며, 표 2에서 MEDIPAM(x2)는 한 번의 데이터증강, 즉 학습용 데이터 양을 2배로 늘린 경우이고, MEDIPAM(x3)는 두 번의 데이터증강 경우를 나타낸다. 데이터 증강의 크기가 커질수록 Dice, IOU 모든 지표가 향상되는 것을 Kvasir 및 CVC-ClinicDB 모두에서 확인할 수 있었으며, 구체적으로 MEDIPAM(x3)는 초해상화만 수행한 MEDIPAM(x1) 대비 USF=2일 때 0.008%, USF=4일 때 0.013% 더 높은 성능을 보여주었다.

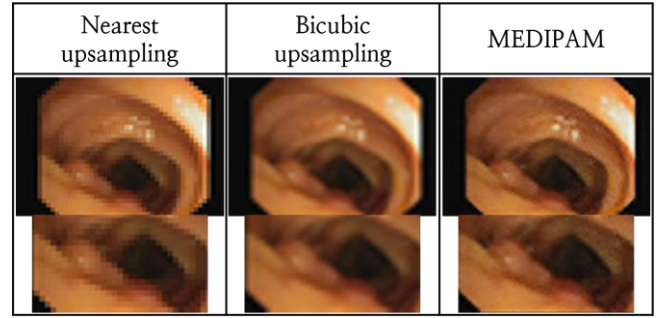


그림 2. 초해상화 기법에 따른 출력 이미지

4. 결 론

본 논문에서는 고해상도 이미지 데이터셋이 부족한 의료 영상 분야에 적합한 초해상화 기법을 연구하였으며, 다량의 훈련데이터 없이 초해상화 기능을 구현할 수 있는 DIP를 활용한 초해상화 및 데이터 증강 기법을 제안하였다. 제안하는 방법은 간단한 DIP 활용을 통해 폴립 영역 분할에 대한 성능 향상을 보여주었으며, 추후 연구에서는 대장 내시경 이미지 이외에도 여러가지 스마트시티 의료 영상 분석에서 활용될 수 있는 DIP 응용 기술을 연구 개발하고자 한다.

참 고 문 헌

- [1] Li, Y., Sixou, B., & Peyrin, F. A review of the deep learning methods for medical images super resolution problems. *Irbm*, 42(2), 120-133. 2021.
- [2] Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, 21(12), 3106-3121. 2019.
- [3] Keys, R. Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, 29(6), 1153-1160. 1981.
- [4] Dong, C., Loy, C. C., He, K., & Tang, X. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307. 2015.
- [5] Lim, B., Son, S., Kim, H., Nah, S., & Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* pp. 136-144. 2017.
- [6] Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., & Fu, Y. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)* pp. 286-301. 2018.
- [7] Ulyanov, D., Vedaldi, A., & Lempitsky, V. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 9446-9454. 2018.
- [8] Ronneberger, O., Fischer, P., & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015: 18th International Conference*, 2015.
- [9] Duc, N. T., Oanh, N. T., Thuy, N. T., Triet, T. M., & Dinh, V. S. ColonFormer: an efficient transformer-based method for colon polyp segmentation. *IEEE Access*, 10, 80575-80586. 2022.