

실제 이미지 초해상도를 위한 학습 난이도 조절 기반 전이학습

조선우 소재웅 조남익

서울대학교 전기정보공학부 뉴미디어통신연구소(INMC)

etoo33_soh90815@ispl.snu.ac.kr nicho@snu.ac.kr

Real Image Super-Resolution based on Easy-to-Hard Transfer-Learning

Cho, Sunwoo Soh, Jae Woong, Cho, Nam Ik

Department of ECE, INMC, Seoul National University

요약

이미지 초해상도는 딥러닝의 발전과 함께 이를 활용하며 눈에 띄는 성능향상을 이루었다. 딥러닝을 기반으로 한 대부분의 이미지 초해상도 연구는 딥러닝 네트워크 모델의 구조에 대한 연구 위주로 진행되어 왔다. 그러나 최근 들어 딥러닝 기반의 이미지 초해상도가 합성된 데이터에 대해서는 높은 성능을 보이지만 실제 데이터에 대해서는 높은 성능을 보이지 못한다는 사실이 주목 받고 있다. 이에 따라 모델 구조를 바꿔 성능을 향상 시키는 것에는 한계가 있어 데이터의 활용이나 학습 방법에 대한 연구의 필요성이 증대되고 있다. 따라서 본 논문은 이미지 초해상도를 위한 난이도 조절 기반 전이학습법(transfer learning)을 제안한다. 제안된 방법에서는 이미지 초해상도를 배울 난이도가 쉬운 낮은 배율부터 순차적으로 전이학습을 진행한다. 이는 이미지 초해상도의 배율이 높아질수록 학습이 어렵기 때문이다. 결과적으로 본 논문에서는 높은 배율의 이미지 초해상도를 진행하기 위해 낮은 배율의 이미지 초해상도, 즉 난이도가 쉬운 학습부터 점진적으로 학습을 진행하였을 때 더욱 빠르고 효과적으로 학습할 수 있음을 보여준다. 제안된 전이학습 방법을 통해 적은 횟수의 업데이트로 학습을 진행하였을 때 일반적인 학습방법 대비 약 0.18 dB의 PSNR 상승을 얻어, RealSR [9] 데이터셋에서 28.56 dB의 성능으로 파라미터 수 대비 높은 성능을 얻을 수 있었다.

1. 서론

이미지 초해상도는 저해상도 이미지를 활용하여 고해상도 이미지를 예측하는 low-level vision 문제 중 하나이다. 이미지 초해상도는 영상 화질의 개선뿐만 아니라 object detection 및 segmentation 등 high-level vision 문제의 성능을 높여주기 위한 전처리 과정으로 많이 사용된다. 따라서, 이미지 초해상도 연구는 매우 중요한 컴퓨터 비전 문제로 여겨져 왔다.

최근 딥러닝의 발전과 함께 이미지 초해상도 문제 또한 이를 활용하여 눈에 띄는 성능향상을 보여주고 있다. SRCNN [1] 은 최초로 이미지 초해상도에 딥러닝을 적용하여 고전적인 방법의 성능을 추월하였다. 뒤 이어 VDSR [2] 과 EDSR [3] 은 이러한 모델 구조를 더욱 깊게 쌓고 잔차 학습방법(residual learning)을 활용한 구조를 도입하여 더욱 눈에 띄는 성능향상을 가져왔다. 또한, 네트워크의 여러 레이어의 특징맵들을 모두 활용하는 residual dense block을 제안한 RDN [4] 은 모델의 파라미터 수를 늘려 성능을 더욱 향상 시켰다. 이 후로, 모델의 채널 간 관심도를 예측하여 좀 더 낮은 파라미터 수로도 더욱 큰 성능을 보여주는 RCAN [5] 과 네트워크를 효율적으로 사용하여 더욱 파라미터 수를 줄인 CARN[6] 이 제안되었다.

기존 연구들은 주로 모델의 구조를 변화시키며 성능향상을 얻어 내는 방법들이다. 이러한 모델들은 합성된 데이터 셋, 즉, noise-free bicubic downsampling의 조건으로 학습이 되었고, 이 조건이 맞을 경우 매우 좋은 성능향상을 기대할 수 있다. 하지만, 이러한 방식으로 학

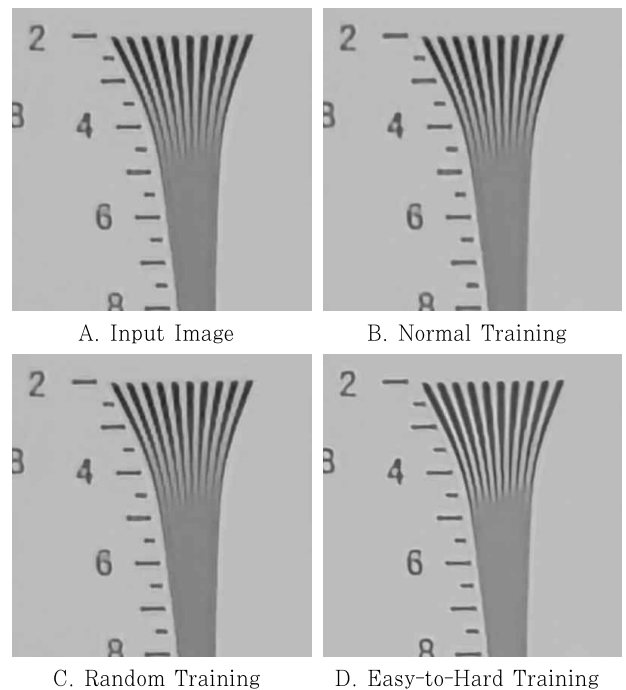


그림1 난이도 조절 학습방법에 대한 결과 이미지.

상좌측부터 A 입력 이미지, B 4배율 데이터만으로 기본적인 학습 방법을 통한 학습결과 이미지, C 2,3,4 배율 데이터를 모두 활용하여 무작위 순서로 학습한 결과 이미지, D 2,3,4배율 데이터를 순차적으로 조절하여 학습한 결과 이미지

습된 이미지 초해상도 방법들은 합성된 데이터와 다른 데이터 분포를 갖고 있는 실제 카메라로 취득된 데이터에 대해서는 높은 성능을 보이지 못한다. 또한, 실제 데이터에서는 합성된 데이터만큼의 많은 양의 데이터쌍을 얻기는 어렵기 때문에 적은 양의 실제 데이터를 활용하는 방법 또는 이를 위한 학습 방법에 대한 연구의 필요성이 나타나고 있다.

이러한 관점에서 한 가지 방법은 적은 양의 실제 데이터를 활용하기 위한 데이터 증대 방법이다. Yoo et al. [7] 는 기존 high-level vision 문제에서 흔히 사용되는 다양한 데이터 증대 방법들을 이미지 초해상도에 적용해보며 이에 대한 성능을 분석해 보았다. 뿐만 아니라 CutBlur 라는 low-level vision 문제 중 특히 이미지 초해상도에서 사용될 수 있는 데이터 증대 방법을 제안하였다. 그러나 위 논문은 적은양의 데이터를 늘리는 것에만 집중하였다. 반면, 이러한 데이터를 이용하여 효과적으로 학습을 진행하는 방법에 대한 연구는 부족하여 해당 방법의 데이터 증대를 사용하였을 때 오랜 시간의 학습이 필요하다. 따라서 본 논문에서는 적은 양의 실제 환경에서 취득한 학습 데이터를 효과적으로 활용하기 위한 전이학습 방법을 제안한다.

전이학습을 위해서 본 논문에서는 이미지 초해상도의 난이도에 초점을 맞추었다. 이미지 초해상도는 저해상도를 고해상도로 키우는 배율에 따라 학습의 난이도가 달라지며, 배율이 낮을수록 학습난이도가 쉽다. 또한, 각 배율의 학습은 같은 문제를 풀고자 하기 때문에 서로 간에 연관성이 있다. 이에 따라 EDSR [3]을 포함하는 많은 논문들에서 배율이 높은 초해상도를 학습할 때 배율이 낮은 초해상도로 학습시킨 네트워크를 기반으로 다시 학습을 시키는 방법을 많이 사용하곤 한다. 이렇게 학습난이도를 쉬운 것부터 어려운 것으로 순차적으로 학습하였을 때 성능이 향상된다는 제안은 high-level vision 문제에서는 과거부터 다양한 방식으로 연구되어 왔으나 [8] low-level vision 문제에서는 많은 연구가 이루어지지 않았다. 따라서 본 논문에서는 부족한 실제 학습 데이터를 보충하기 위한 전이학습 방법을 제안한다. 구체적으로, 특정 배율에서 부족한 학습 데이터를 다른 배율의 학습 데이터를 활용함으로써 보충할 뿐 아니라 이 데이터를 효과적으로 활용하기 위해 데이터의 배율을 순차적으로 증가시키며 학습을 진행한다. 그림 1을 보면 이러한 학습 방법을 활용하였을 때 단순히 데이터를 증가시켰을 때에 비해 월등한 성능향상을 보여줄 수 있다. low-level vision 문제에서도 난이도가 쉬운 학습부터 점진적으로 학습을 진행하였을 때 더욱 효과적인 학습을 할 수 있음을 보여주었다. 본 논문에서 제안된 방법을 통해 약 0.18 dB PSNR 의 상승을 얻을 수 있었을 뿐 아니라 1.4 M의 파라미터로 8000번의 적은 반복에서 실제 데이터셋인 RealSR [9] 데이터셋에서 28.56 dB 라는 높은 성능을 얻을 수 있었다.

2. 실험 방법

2.1 네트워크 구조

본 논문에서 사용한 전체적인 네트워크 구조는 그림 2에 잘 나타나 있다. 이때 그림 2에서 사용한 residual block (Resblock)의 구조는 그림 3과 같다. 네트워크를 효율적으로 구성하기 위해 네트워크 앞단에 desubpixel convolution 을 활용하여 3채널 이미지를 12채널로 늘리는 대신 입력 이미지의 해상도를 낮추었다. 이어서 잔차학습 방법을 활용한 그림 3와 같은 Resblock 구조를 16개 쌓았다. 마지막으로

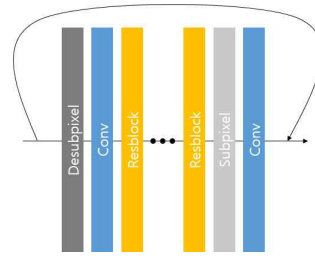


그림 2 네트워크 구조

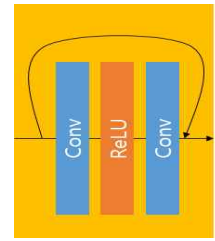


그림 3 Resblock 구조

subpixel convolution 을 활용하여 특징맵의 해상도를 다시 입력 이미지의 해상도와 동일하게 키운 후 skip connection을 활용하여 이미지의 잔차를 학습할 수 있도록 네트워크를 설계하였다.

2.2 학습 방법

본 논문에서 제안하는 학습 난이도 조절에 따른 전이학습에서 학습의 난이도는 이미지 초해상도의 배율에 비례하는 것을 이용하였다. 이미지 초해상도는 배율이 높을수록 주어진 픽셀 대비 예측해야하는 픽셀의 개수가 많아지므로, 또한, 더 넓은 고주파 영역 대를 예측해야하므로 더 난이도가 높은 문제라고 할 수 있다. 따라서 배율이 낮을수록 쉬운 난이도에 속하고 배율이 높을수록 어려운 난이도에 속한다. 따라서, 본 방법은 이러한 특징을 기반으로 특정 배율에 대한 학습을 진행할 때 다른 배율의 데이터를 활용한다. 구체적으로는 낮은 배율부터 순차적으로 학습을 진행한다. 예를 들어 4배율을 학습하기 위해서는 가장 먼저 2배율 데이터를 활용하여 2배율에 대한 학습을 먼저 진행한 후 학습이 완료되면 이렇게 학습된 파라미터를 기점으로 하여 3배율 데이터를 활용하여 3배율에 대한 학습을 진행한다. 이렇게 3배율 또한 학습이 완료되면 이렇게 학습된 파라미터를 시점으로 하여 최종적으로 4배율 데이터를 활용하여 4배율 학습을 진행한다.

이때 각 배율의 학습 단계에서는 주어진 데이터의 정답과 예측된 고해상도 이미지 간의 평균 제곱 오차(Mean Squared Error)를 측정하여 이를 줄이는 방향으로 학습을 진행하였다.

3. 실험 결과

실험은 실제 카메라로 취득한 저해상도 이미지와 고해상도 이미지가 쌍으로 존재하는 RealSR [9] 데이터셋을 이용하여 학습 및 평가를 진행하였다. RealSR 데이터셋 [9] 중 Canon dataset과 Nikon dataset의 모든 배율 이미지 (153장의 2배율 이미지, 204장의 3배율 이미지, 148장의 4배율 이미지)를 활용하여 4배율 초해상도를 할 수 있는 모델의 학습을 진행하였다. 학습은 Adam optimizer를 활용하였고 learning rate는 2배율과 3배율에서는 0.0001, 4배율에서는 0.00005를 사용하였다. 배율에 따른 단계적 학습에서 2배율 데이터에 대해서는 4000번의 반복을 통해 학습을 진행했고 3배율에 대해서는 2000번의 반복, 마지막으로 4배율에 대해서는 2400번의 반복만큼 학습을 진행하였다.

표 1 다른 모델들과의 성능 비교

RealSR dataset 중 4배율에 대한 성능측정 결과. 이때 초기 8,000번의 iteration에서의 성능을 측정하기 위해 공개된 소스코드로 직접 실험을 진행하였다.

	CARN	CARN+CutBlur	제안된 방법
PSNR	28.40	28.44	28.56
파라미터 수	1.14M	1.14M	1.4M

3.1 다른 모델들과 성능 비교

제안한 방법을 이용한 학습을 통하여 표 1과 같은 성능을 얻을 수 있었다. 이 때 다른 모델들에 비해서 적은 학습 시간동안에도 효과적으로 학습이 진행됨을 보여주기 위해 8000번의 반복에서의 성능을 측정하였다. 제안된 방법은 비슷한 파라미터 수를 사용하였지만, 모델 구조를 효율적으로 변화시킨 것(CARN)보다 추가적인 데이터를 활용하여 학습을 효과적으로 진행한 방식이 더 높은 성능 향상을 보였다. 이를 통해 데이터가 적은 상황에서는 모델의 구조뿐만 아니라 다양한 데이터를 활용한 학습방식 또한 중요함을 알 수 있다. 또한 비슷한 크기의 모델에 CutBlur의 데이터 증대 방식을 활용한 학습방법에 비해서도 적은 학습 횟수에서 더 높은 성능을 보여준다.

표 2 다른 학습방법과의 성능비교

RealSR dataset 중 4배율에 대한 성능 측정 결과 이다. 이때 각 학습방법은 모두 같은 모델에서 진행하였고, 모두 무작위의 파라미터 초기값에서 학습을 시작하였다.

	(1) 4배율 데이터로만 학습	(2) 2,3,4 배율 데이터를 무작위 순서로 학습에 활용	(3) 2,3,4배율 데이터를 순차적으로 학습에 활용
PSNR	28.38	28.48	28.56

3.2 다른 학습방법과 성능비교

제안된 학습방법의 효과성을 입증하기 위해서 학습방법의 각 요소를 없앤 다른 학습방법과의 성능비교 실험을 진행하였다. 총 2개의 추가 실험을 진행하였는데, 우선 가장 기본이 되는 실험은 기존의 일반적인 학습방식인 4배율의 데이터로만 학습을 진행하는 것이다. 두 번째 학습 방식은 2, 3, 4 배율 데이터를 모두 활용하였지만, 난이도에 기반하여 순차적인 전이학습을 진행하지 않고 2, 3, 4배율의 데이터를 무작위로 사용하여 학습을 진행한 것이다. 이 두 추가 모델에 대한 실험결과는 위의 표 2와 같다. 표를 확인하였을 때 (2)번의 실험의 결과가 (1)번의 결과에 대비하여 추가적인 데이터를 사용하였음에도 불구하고 이러한 데이터를 효과적으로 사용하지 못하여 성능향상에는 큰 영향을 미치지 못하는 것을 알 수 있다. 결론적으로 데이터의 절대적인 양이 적을 경우, 추가적인 데이터를 사용할 때에는 효과적으로 사용을 위해 학습방법 또한 함께 고려해야 함을 알 수 있다.

그림 4는 (1), (2), (3) 번의 각 학습과정에서 각 학습 반복 수에 따른

4배율 test dataset 의 PSNR 값의 변화를 그린 것이다. 이때 파란색 그래프가 (1)번 학습과정, 주황색 그래프가 (2)번 학습과정, 노란색 그래프가 제안된 방법, 즉 (3)번의 학습과정이다. (1)번 학습방식은 테스트 환경과 동일한 데이터만을 사용하여 학습은 빠르지만 금방 모델이 포화되는 것을 확인할 수 있다. 또한, 여러 데이터를 모두 활용한 (2)번 학습과정의 경우 학습 속도는 느리지만 최종적으로는 (1)번과 비슷한 성능을 보여줄 수 있다. 그러나 본 논문에서 제안하는 (3)번 학습 방법의 경우 초기에는 다른 배율의 데이터를 통해 학습을 진행해서 학습 환경과 테스트 환경의 불일치 문제로 성능이 좋지 않다가 학습과 테스트 환경이 비슷해지는 3배율 학습 과정에서부터 눈에 띄는 성능향상을 보여주기 시작한다. 결론적으로 최종 성능에서는 다른 두 학습방법에 비해 높은 성능을 보여준다.

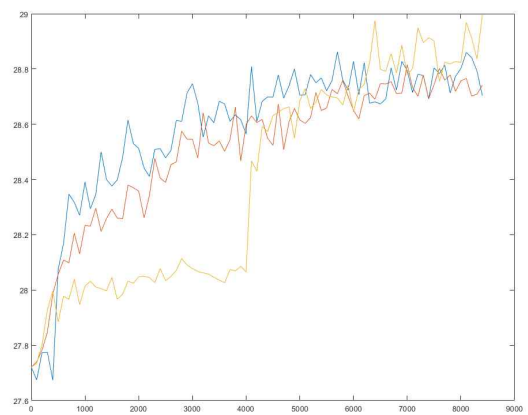


그림 4 다른 학습 방법들의 학습과정 중 성능 변화 그래프

이때 가로축은 학습 반복 수이고 세로축은 PSNR이다. RealSR dataset 중 Canon dataset에서 성능을 측정하였다.

4. 결론

본 논문에서는 난이도 조절에 기반한 전이학습방법을 활용하여 학습 이미지 데이터가 적은 실제 환경의 이미지에 대한 학습을 효과적으로 하는 방법을 제안하였다. 실험 결과를 통해 본 논문에서 제안한 학습방법을 활용하였을 때 같은 데이터셋을 활용하더라도 데이터를 효과적으로 활용하여 눈에 띄는 성능향상을 보여주었다. 본 연구는 적은 실제 데이터에 대해 적용 가능하지만 실제 데이터가 모든 스케일에 대해 구비되어 있어야 한다는 한계를 갖고 있어 추후 이를 보완하는 방법에 대한 연구가 기대할 수 있을 것으로 보인다.

감사의 글

이 논문은 2020년도 BK21 플러스 사업에 의하여 지원되었음.

참고 문헌

[1] Dong Chao, et al. "Learning a deep convolutional

- network for image super-resolution." *In ECCV 2014*
- [2] Jiwon Kim et al. "Accurate image super-resolution using very deep convolutional networks." *In CVPR 2016*
- [3] Bee Lim et al "Enhanced Deep Residual Networks for Single Image Super-Resolution". *In CVPR 2017 Workshops.*
- [4] Yulun Zhang, et al. "Residual dense network for image super-resolution." *In CVPR 2018*
- [5] Yulun Zhang, et al. "Image super-resolution using very deep residual channel attention networks." *In ECCV 2018*
- [6] Namhyuk Ahn, et al. "Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network" *In ECCV 2018*
- [7] Jaejun Yoo, et al. "Rethinking Data Augmentation for Image Super-resolution: A Comprehensive Analysis and a New Strategy" *In CVPR 2020*
- [8] Yoshua Bengio et al. "Curriculum Learning" *In ICML 2009*
- [0] Jianrui Cai, et al. "Toward Real-World Single Image Super-Resolution : A New Benchmark and A New Model" *In ICCV 2019*