

## 훈련 데이터셋의 조절을 통한 딥러닝 기반 Super-Resolution 의 성능 향상

권기택, 서영호

광운대학교

[kitaek0622@gmail.com](mailto:kitaek0622@gmail.com), [yhseo@kw.ac.kr](mailto:yhseo@kw.ac.kr)

### Performance Enhancement of Deep Learning-based Super-Resolution by Adjustment of Training Dataset

Ki-Taek Kwon, Young-Ho Seo

Kwangwoon University

#### 요 약

본 논문에서는 CAR(content adaptive resampler)로 축소된 저해상도 이미지를 직접 다른 모델에 여러가지 방식으로 훈련을 시켜 성능을 개선시키고자 하였다. 본 논문에서는 단일 영상 super resolution 에 관하여 여러 기술이 존재하는 상황에 더 나은 기술을 테스트하려 하고 그를 위해 과거의 모델들에 대한 이해가 필요하여 이를 구현하였다. 현재 가장 뛰어난 성능을 보이고 있는 모델 중의 하나인 CAR 에서 복원 전 이미지를 사용하여 훈련을 시키면 더 나은 성능의 모델을 만들 수 있을 것이라고 가정하고 다양한 훈련을 통해 성능을 개선시키고자 하였다.

#### 1. 서론

최근 몇 년간 super-resolution 기술이 많은 관심을 받아왔다. Super-resolution 이란 Deep-learning 기술을 이용하여 저해상도의 이미지를 고해상도의 이미지로 변환시키는 기술로 CCTV 혹은 TV 등 다방면에서 사용이 가능하며 뛰어난 효과를 가질 것으로 예상된다. 실제로 현재 많은 프로그램들이 만들어져 사용되고 있다. 사람들은 저해상도의 이미지보다 고해상도의 이미지가 이미지를 표현하는 측면에서 더 뚜렷하고 자세하여 실제처럼 보이기 때문에 고해상도를 선호한다. 또한 확대가 필요한 CCTV 나 TV 에서 이미지의 픽셀이 계단처럼 보이는 현상. Aliasing 을 방지하고 TV 자체적으로 super-resolution 기술을 갖게 된다면 저해상도의 영상을 송신을 받고 이를 고해상도로 복원을 시켜 전송효율 측면에서 더 뛰어난 효과를 얻을 것이라고 기대된다. 많은 super-resolution 기술

중에 몇 가지를 선정하여 그 기술들의 구조를 분석하고 구현하여 실제로 어떤 차이가 있는지 명확히 알 수 있도록 하고 이를 현재 가장 큰 성능을 갖는 CAR 과 비교해보려고 한다. CAR 의 경우에는 자체적으로 이미지를 저해상도로 낮추고 이를 다시 고해상도로 복원하는 형식의 모델이기 때문에 들어간 이미지의 픽셀이 출력되는 이미지의 픽셀이 된다. 이는 확대의 측면으로 볼 수 없고 복원의 측면으로 봐야 하기 때문에 확대하는 모델로서 사용할 수 없다.

#### 2. Super-resolution 모델

Super-resolution 은 많은 모델이 있지만 모델의 개선 과정을 보기 쉽도록 모델을 선정하였다. 첫번째 모델은 SRCNN 으로 딥러닝을 super-resolution 에 적용한 초기의 방식으로 비교적 간단한 구조로 되어있다[1]. 두번째 모델은

FSRCNN 으로 SRCNN 이 속도 측에서 느린 것이 입력 이미지의 픽셀 값이 큰 것이 원인이라고 생각하여 입력 이미지보다 확대된 이미지가 출력이 되는 구조이다[2]. 학습속도를 측정하였을 경우 SRCNN 은 step 당 30 초인 반면에 FSRCNN 은 7 초의 학습속도를 보였다. 세번째 모델은 ESPCN 으로 이미지 확대의 방식을 Deconvolution 을 사용하지 않고 sub-pixel convolution 이라는 방식을 도입하였다[3]. 네번째 모델은 EDSR 이다. EDSR 은 전의 layer 에 의한 값만 입력으로 받는 전 모델들의 방식과는 다르게 첫 입력도 같이 입력으로 받는 구조인 residual unit 을 사용한다 [4]. 다섯 번째 모델은 AWSRN 으로 파라미터의 수가 적은 경량형 모델이다[5].

마지막 모델은 CAR 로 그림 1 에 구조를 나타내었다. CAR(Content Adaptive Resample)는 세 가지 부분으로 구성이 되어있다. 이미지를 줄이고 픽셀 값을 특징하는 커널 생성 네트워크, 줄인 픽셀 값으로 저해상도 이미지를 형성하는 downsampler 네트워크와 이미지를 다시 확대하는 upscale 네트워크로 나뉘어 진다. 이때 upscale part 는 EDSR 을 사용하였다[6].

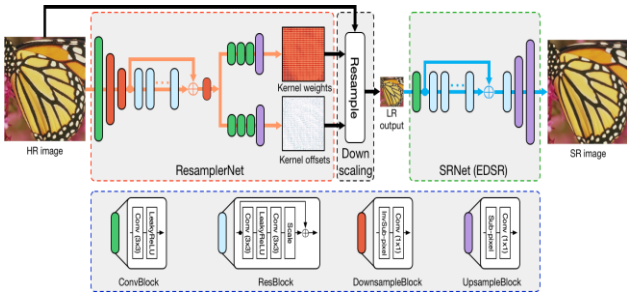


그림 1. CAR 모델의 구조

### 3. 훈련데이터에 의한 SR 성능 분석

6 개의 모델을 전부 학습하였고 Set5 모델로 테스트를 수행 하였다. 전부 2 배 확대를 기준으로 하였고 사용한 학습모델은 Div2k 이고 optimizer 는 Adam 을 사용하였다. 표 1 에 결과를 정리하였다.

표 1. Super-resolution 모델들의 성능 평가 결과

	SET5_PSNR	SET5_SSIM
SRCNN	25.842188	0.918666
FSRCNN	19.749450	0.704651
ESPCN	19.961820	0.718296
EDSR	20.796150	0.724138
AWSRN	20.090850	0.670540
CAR	38.95	0.9648

### 4. CAR 를 이용한 학습과정

표 1 에서 CAR 가 가장 뛰어난 정확도를 보였다. CAR 의 정확도의 원인은 복원 성능이 좋다는 측면도 있지만 복원이 잘되도록 저해상도 영상을 생성한다는 측면도 있다. Downsample 의 과정에서 불필요한 정보를 제거하였고, 복원에 필요한 값만 남은 이미지라고 예측할 수 있다. 이러한 효과를 관찰하기 위해 Div2k 와 CAR 의 downsample 이된 Div2k 로 학습한 후에 해상도를 증가시키는 실험을 수행하였다.

Set5 각각의 사진의 정확도를 비교하는 과정에서 일부의 영상은 기존 그대로가 좋은 결과가 나오고 일부의 영상은 downsample 된 div2k 로 학습된 모델에서 성능이 우수한 것을 확인하였다. 따라서 두 종류 모두의 정확도 향상을 위해서 기존의 div2k 와 CAR 로 downsample 한 div2k 를 50 : 50 으로 혼합하여 학습을 시도해 보았다. 이러한 결과들을 표 2 에 정리하였다. 또한 Set5 전체로 비교하면, 표 3 와 같은 결과를 얻을 수 있다.

표 2. 저해상도의 생성 방식에 따른 SR 효과 비교

PSNR by Set5_butterfly	X2	X4
div2k 학습	30.80211	29.39421
Downsample div2k 학습	30.6912	29.3934
5:5 비율	30.87789	29.461382

표 3. Set5 에 대한 모델 정확도의 비교

PSNR	Set5_x2	Set5_x4
Org_div2k	32.92967	31.22771
Car_div2k	32.94379	31.23861
5:5_div2k	33.04164	31.29872

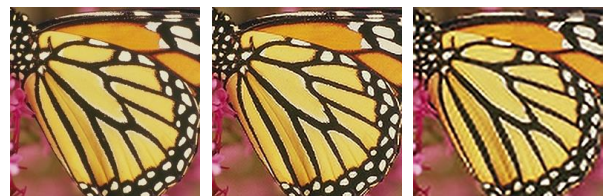


그림 2. 5:5 의 비율로 데이터셋을 구성한 경우의 결과 (a) 2 배 확대, (b) 4 배 확대

### 5. 결론

여러 모델들의 구조를 구현을 통해 이해하였고 모델의

발전 볼 수 있었다. 처음에는 속도를 위하여 마지막에 이미지를 확대하는 방식을 사용하였지만 이는 낮은 정확성으로 이어졌고 다시 정확도 향상을 위하여 파라미터와 연산량을 늘려 모델의 정확도를 높였지만 그 결과는 다시 무거워진 모델로 실제로 사용하기 어려운 모델이 되었다. 모델 학습과정에서 learning rate 나 optimize 등의 최적화를 통해 성능을 향상시킬 수도 있지만, 많은 시간과 비용이 소모된다. 이와 같은 이유로 인해 우리는 데이터셋의 측면에서, 데이터 세트의 적절한 비율 조정을 통해 얼마든지 성능을 향상시킬 수 있는 가능성을 확인하였다.

## 6. 감사의 글

이 논문은 2021 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2018R1D1A1B07043220).

## 7. 참고문헌

- [1] Dong, Chao, et al. "Image super-resolution using deep convolutional networks." *IEEE transactions on pattern analysis and machine intelligence* 38.2 (2015): 295-307.
- [2] Dong, Chao, Chen Change Loy, and Xiaoou Tang. "Accelerating the super-resolution convolutional neural network." *European conference on computer vision*. Springer, Cham, 2016.
- [3] Shi, Wenzhe, et al. "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [4] Lim, Bee, et al. "Enhanced deep residual networks for single image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
- [5] Chaofeng Wang, Zheng Li, Jun Shi. "Lightweight Image Super-Resolution with Adaptive Weighted

Learning Network." *Computer Vision and Pattern Recognition*, Shanghai University, 2019

- [6] Wanjie Sun, Zhenzhong Chen "Learned Image Downscaling for Upscaling using Content Adaptive Resampler" *School of Remote Sensing and Information Engineering*, Wuhan University, 2019.