

딥러닝을 이용한 low dose image를 high dose image로 변환하여 결과 비교

성명 : 2021210088 허지혜

1) Image Super Resolution(SR)

Image Super Resolution은 저해상도 (Low Resolution) 이미지를 고해상도 (High Resolution) 이미지로 변환시키는 문제를 의미한다. 한 장의 저해상도 영상을 가지고 고해상도 버전을 만들어 내는 기술을 SISR (Single Image Super Resolution) 이라고 하고, 여러 장의 저해상도 영상을 가지고 고해상도 버전을 만들어 내는 기술은 MISR (Multi Image Super Resolution) 이다. 보통 MISR과 SISR를 비교했을 때 SISR이 많이 쓰이는데 이유는 인풋 개수가 하나이니 훨씬 효율적이기 때문이다.

SISR에서 나타나는 문제점은 ill-Posed 문제이다. Super Resolution은 저해상도 이미지를 고해상도로 복원을 해야 하는데, 복원해야 하는 타겟인 고해상도의 이미지가 정답이 여러 개 존재할 수 있다. 어떻게 만들어도 정답이 되는 즉, 정답이 유일하지 않다는 의미이다. 복원된 고해상도가 잘 복원되었는지 확인하기 위해선 몇 가지 기준이 있지만 그건 뒤에서 다시 설명할 것이다.

SISR에 대해서 간단하게 설명하자면, 저해상도 이미지를 y 라고 했을 때 다음과 같은 수식을 따른다.

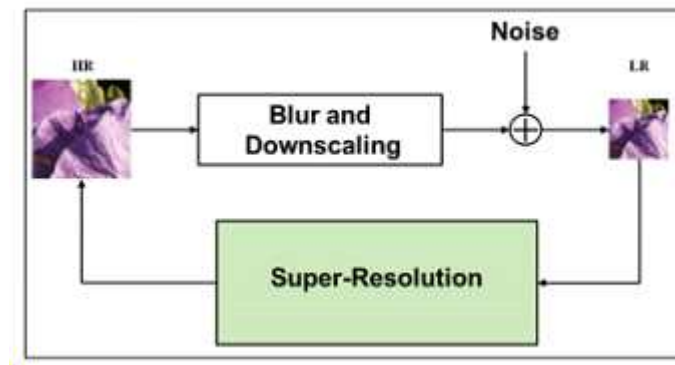
$$y = (x \otimes k) \downarrow_s + n$$

$x \otimes k$: 고해상도 이미지 x , Blur kernel k

down s : 이미지 해상도를 다운시키는 비율 s

n : 독립적인 noise

위를 그림화하면 다음과 같이 나타낼 수 있다.



일반적으로 HR(High Resolution) 이미지를 GT(Ground Truth)로 정의하고 LR(Low Resolution)으로 만들기 위해 blurring, down sampling, noise 등을 임의로 거쳐 저해상도 이미지로 만든다. 그 뒤 어떤 방법을 통해 저해상도 이미지를 GT로 복원시키도록 모델을 학습시키는 구조를 따른다. 그치만 저해상도 이미지를 만들 때 사용한 기법이 무엇이었는지에 따라 Super Resolution 성능이 달라질 수 있다.

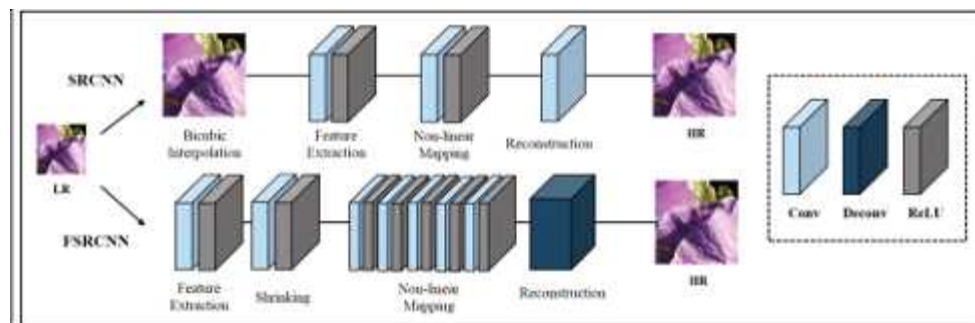
SR의 접근 방법은 크게 3가지가 존재한다.

1. Interpolation-based method
2. Reconstructed-based method
3. Deep learning-based method

저해상도 이미지라는 것은 pixel 수가 작다는 의미이기 때문에 1번 방법을 이용한 연구가 먼저 많이 이루어졌다고 한다. 이미지 변환 라이브러리인 CV를 이용하여 이미지의 크기를 바꿔본 경험은 1번 방법에 속한다. Interpolation(보간법) 옵션을 다르게 사용하면 이미지의 품질이 달라진다. Bicubic, Bilinear, Nearest Neighbor 등 다양한 옵션이 있으면 일반적으로 이미지의 해상도를 키워주는 경우엔 Bicubic, Bilinear, Lanczos 보간법을 사용한다.

하지만 위 방법은 이미지의 크기를 크게 만들어줄 뿐, 디테일한 부분은 신경쓰지 못한 경향을 보인다. 2번 방법은 저해상도 이미지를 고해상도 이미지를 만드는 방법 자체를 연구하는 방향으로 시도하였는데 빠르고 좋은 결과를 가져왔다. 이를 example-based SISR이라고 부르기도 한다. 또 마지막 방법은 Deep Learning-based로 2번보다 매우 우월한 결과를 가져올 수 있다. 2번, 3번 방법이 자주 등장하게 되었다.

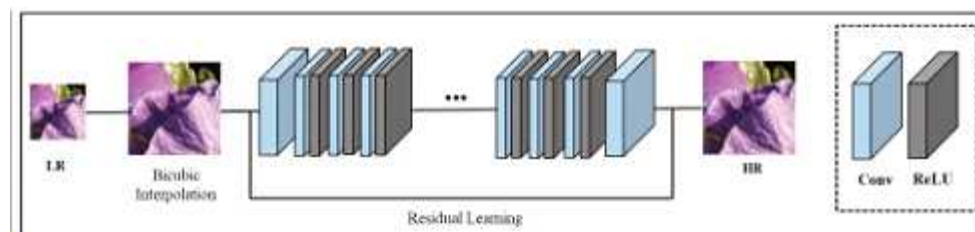
2) 딥러닝을 이용한 Super Resolution



- SRCNN

3계층의 FCN(Fully Convolutinal Network) 구조를 가진 네트워크에 bicubic Interpolation으로 확대한 이미지의 화질을 개선하는 알고리즘이다.

여기 알고리즘보다 조금 더 발전한게 FSRCNN이다.



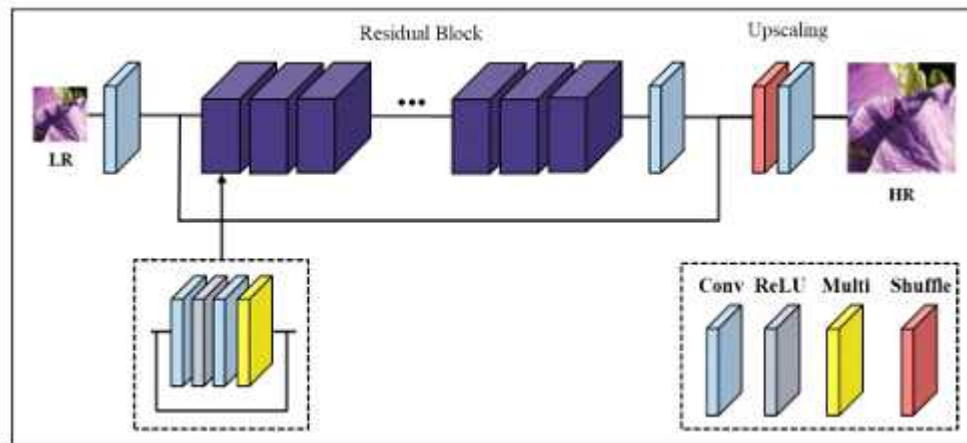
- VDSR(Very Deep Super Resolution)

딥러닝 학습을 이용한 영상초해상도 알고리즘

VGG 기반의 20 layer 깊이의 CNN 구조를 처음 사용

input image를 최종 output에 더해주는 Residual 학습 개념 도입

높은 learning rate를 사용해 정확도를 획기적으로 높임



- EDSR(Enhanced Deep Super Resolution)

32 layer 이상을 사용하고 channel 수도 기존의 네트워크에 대비해 4배 이상 늘려 파라미터 수도 그에 비례하여 매우 많이 증가하였다. 네트워크를 보다 안정적으로 학습하기 위해 네트워크를 Residual Block별로 나누고 각각 Skip connection을 사용해 연결하여 필터 파라미터들을 더욱 쉽게 최적화되도록 네트워크를 설계하였다.

또한 일정 상수 값을 CNN 이후에 곱해주는 Multi layer를 추가하였다.

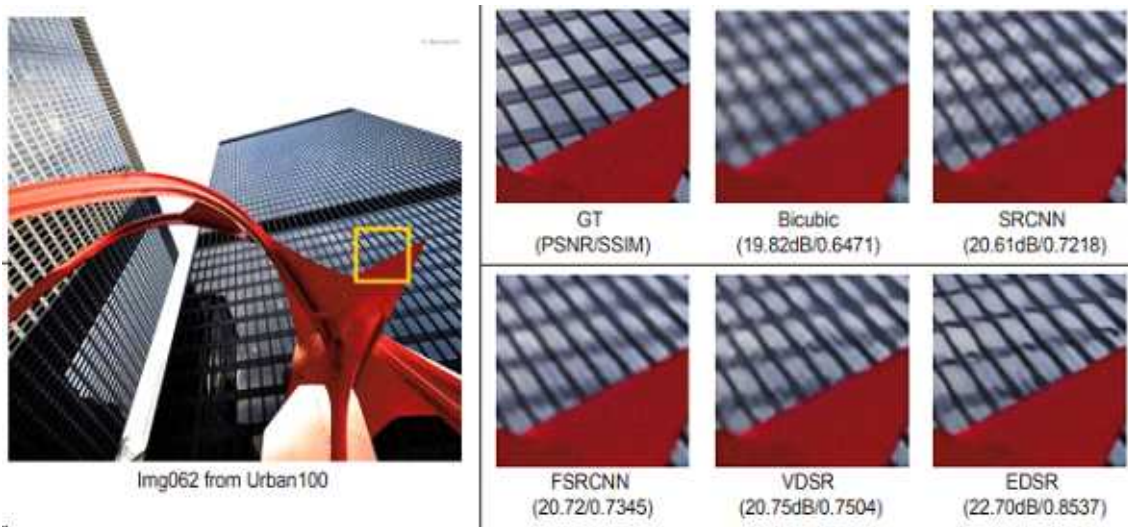
3) 성능 지표

정량적 평가는 지표를 통해 나온 수치이고 정성적 평가는 실제 이미지를 확인 하였을 때 정말 선명하게 복원이 되었는지 확인하는 평가이다. 영상의 화질을 복원했는지 정량화 하기 어렵기 때문에 성능 지표 또한 까다롭다. 대부분의 논문에서는 성능 지표로 PSNR과 SSIM(Structural SIMilarity)을 사용한다.

PSNR : 압축된 영상의 화질에 대한 손실 정보를 평가한다. 그러나 PSNR 값이 유사하더라도 화질의 차이가 있을 수 있다는 단점을 가지고 있다. 또한 PSNR은 비교 지표이기 때문에 절대적인 수치로 받아들이기 보다는 정상적으로 고려하여 평가하는게 좋다.

SSIM : 인간이 이미지를 어떻게 보는지를 착안하여 만들어진 지표로 이미지의 구조 정보가 얼마나 유사한가를 평가해주는 지표이다.

하지만 PSNR, SSIM이 높다고 해서 사람 눈에 보기 좋은 것은 아니다. 다른 논문에서는 MOS(Mean Opinion Score) 지표도 사용하기도 한다.



*) Reference

[Single Image Super Resolution using Deep Learning Overview] :

<https://hoya012.github.io/blog/Single-Image-Super-Resolution-Overview/>

<http://www.kibme.org/resources/journal/20200504094149078.pdf>

<http://dmqm.korea.ac.kr/activity/seminar/318>

<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=rhukjin&logNo=222064619296>