

I. 서론

FHD가 시장의 주요 사용 해상도였던 과거와 달리, 최근 들어 QHD, UHD와 같은 고해상도 디스플레이가 대중화되면서, 저해상도(Low Resolution, LR) 이미지에서 고해상도(High Resolution, HR) 이미지를 복구하기 위한 기술인 초해상화(Super Resolution, SR) 알고리즘이 주목받기 시작하였다.

초해상화 알고리즘은 컴퓨터 비전 및 영상처리에서 주요한 기술 중 하나이며, 의료 영상, 감시 및 보안과 같이 다양한 분야에서 응용되고 있다.

하지만 단일 LR 영상에 해당하는 여러 개의 HR 영상이 항상 존재하므로 이 문제는 본질적으로 잘못 되어있다. 이를 부적절하게 정립된 역 문제(ill-posed inverse problem)로 정의한다.

딥러닝 기반 방법들이 이용되기 전에는 일반적으로 예측 기반 방법, 엣지 기반 방법, 통계 기반 방법, 패치 기반 방법 및 희소 표현 방법 등의 방법이 제안되어왔다. 하지만 최근 몇 년간 GPU의 성능이 급속도로 발전하면서 딥러닝을 활용한 기법들 역시 빠르게 개발되었다. 발전된 딥러닝 기법들은 다양한 초해상화 벤치마크에서 기존 기법 대비 최첨단의 성능을 달성하고 있는 경우가 많다.

딥러닝 기반 초해상화 기법들은 초기 CNN 기반한 SRCNN부터 생성적 적대 신경망에 기

반한 SRGAN까지, 다양한 방법들이 적용되어왔다. 본 보고서에서는 딥러닝 기반 초해상화 기법들을 3가지의 방법으로 분류하여 각 방법에 대하여 설명한 뒤, 앞서 언급한 두가지 네트워크와 그 중간에 제안된 기법들을 망라하여 초해상화 알고리즘의 발전 동향을 다룰 것이다.

II. 초해상화 기술 동향

1. 초해상화 방법

가. 사전 업샘플링 초해상화

기본적으로 SR 문제는 저차원 공간에서 고차원 공간으로의 매핑을 학습하는 것으로, 매우 어려운 문제에 해당한다. 이를 해결하기

위해 기존의 업샘플링 방법(e.g., 바이큐빅 알고리즘)을 활용하여 고해상도 영상을 획득한 다음, 이를 심층 신경망을 사용하여 refine 하는 것이 간단한 해결책으로 제안되었다. 이러한 방식에 기반한 모델들은 일반적으로 LR 영상을 기존 알고리즘을 사용하여 원하는 크기로 업샘플링된 HR 영상을 획득하고, 이를 다시 CNN 을 이용하여 세부적인 정보를 복원하는 방식을 이용한다. 가장 어려운 부분이 업샘플링 작업이었던만큼, CNN 은 단순히 영상을 다듬는 정도의 작업을 수행하게 되므로 학습 난이도가 현저하게 낮아진다. 또한 이러한 방식이 가지는 장점은 임의의 크기와 스케일링 팩터를 입력으로 하지만, 단일 스케일링 팩터에 기반한 초해상화 모델들과 유사한 성능을 보여줄 수 있다는 것에 있다.

따라서 초해상화 초기 가장 인기있는 방식중 하나가 되었다.

그러나 이 방식은 기존 알고리즘에서 발생하는 문제였던 노이즈 증폭 및 블러링과 같은 문제를 여전히 갖고 있었으며, 대부분의 계산처리는 HR 영상과 같은 고차원 공간에서 진행된다. 따라서 타 방식 대비 시간 복잡도는 필연적으로 높은 방식이다.

나. 사후 업샘플링 초해상화

이 방식은 사전 업샘플링 초해상화에서 이용되었던 기존 알고리즘을 이용한 업샘플링 이후 CNN 을 이용한 세부정보 복원에서, 기존 알고리즘까지 신경망을 이용하여 대체하는 방식이다. 이를 통해 계산 효율성을 개선하고, 기존 업샘플링 방식이 end-to-end 방식으로 학습 가능한 레이어들로 대체된다. 따라서 CNN 에 입력되는 영상은 HR 영상이 아닌 LR 영상이 되며, 학습 가능한 업샘플링 레이어가 네트워크 끝에 적용되는 것이 일반적이다. 그러므로 매우 큰 계산비용을 필요로 하는 feature extraction 과정은 저차원 공간에서만 발생하고, 네트워크의 끝에서 해상도가 높아지므로, 전체적인 계산과 공간 복잡성이 사전 업샘플링 초해상화 방식 대비 크게 줄어든다.

다. 점진적 업샘플링 초해상화

사후 업샘플링 초해상화 방식은 사전 업샘플링 초해상화 대비 계산과 공간 복잡성을 크게 낮췄지만, 여전히 단점은 존재했다. 첫번째로 업샘플링은 한 단계로만 수행되므로, 큰 스케일링 팩터에 대한 학습 어려움이 크게 증가한다. 또한 각 스케일링 팩터에 대해 개별적으로 초해상화 모델을 학습시켜야 하며, 다중 스케일링 팩터를 필요로 하는 경우에는 대처할 수 없다. 이를 해결하기 위해 제안된 방식이 바로 점진적 업샘플링 SR 이다. 이

방식은 다단계 CNN 에 기반하여 점진적으로 HR 영상을 복원한다. 각 단계별로 영상은 더 높은 해상도로 업샘플링 - 정제 과정을 거치게 된다.

2. Super-Resolution 모델

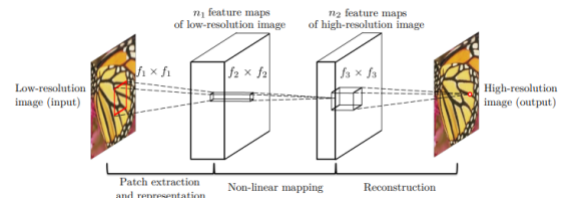


그림 1. SRCNN의 모델 구조[1]

가. SRCNN

Dong 에 의해 제안된 SRCNN[1]은 사전 업샘플링 초해상화 분야에 최초로 CNN 을 적용한 연구이다. 이전의 사전 업샘플링 초해상화 연구들의 경우, 여러 단계의 알고리즘을 통해 사전 업샘플링 초해상화 작업이 수행되었다. 반면 SRCNN 은 최초의 end-to-end 방식으로, 모든 연산작업이 하나의 네트워크에서 수행되었다. 이 모델은 전형적인 사전 업샘플링 초해상화 방식으로, Bicubic 알고리즘을 통해 LR 영상을 업스케일링 한 뒤, 이를 CNN 네트워크를 통해 보완하는 방식이다.

SRCNN 은 총 3 개의 FCN(Fully Connected Network) 레이어 구조를 가지고 있으며. 각각의 레이어에서 수행되는 역할에 대해서 분석하였다. 첫번째 레이어는 패치 추출 과정으로, 입력된 업샘플링된 LR 영상을 패치단위로 분할하여 각 패치에 대해 특징을 추출하는 과정이다. 두번째 레이어는 추출된 특징들을 입력으로 하여 비선형적인 맵핑을 통해 또다른 특징으로 변환한다. 세번째 레이어는 두번째 레이어에서 맵핑된 특징들을 이용하여 최종 HR 영상을 생성하는 과정이다.

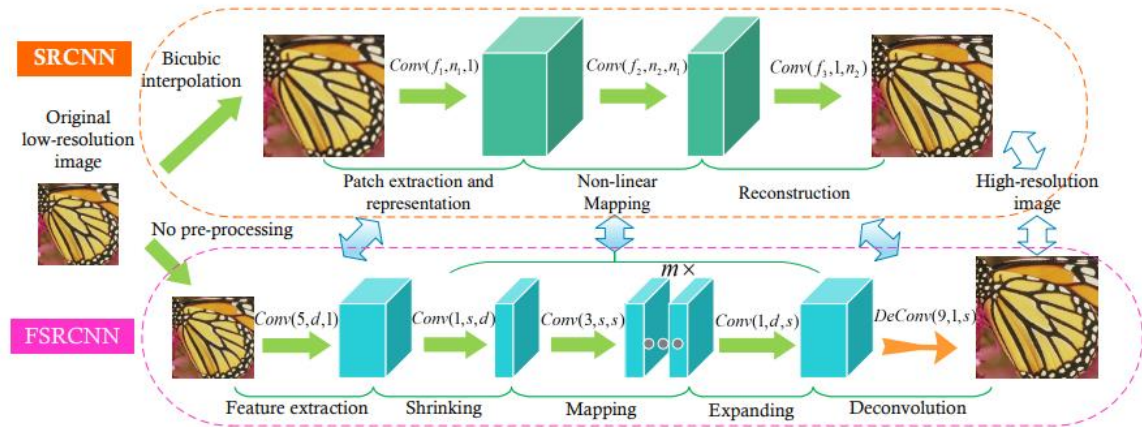


그림 2. SRCNN 및 FSRCNN의 모델 구조[2]

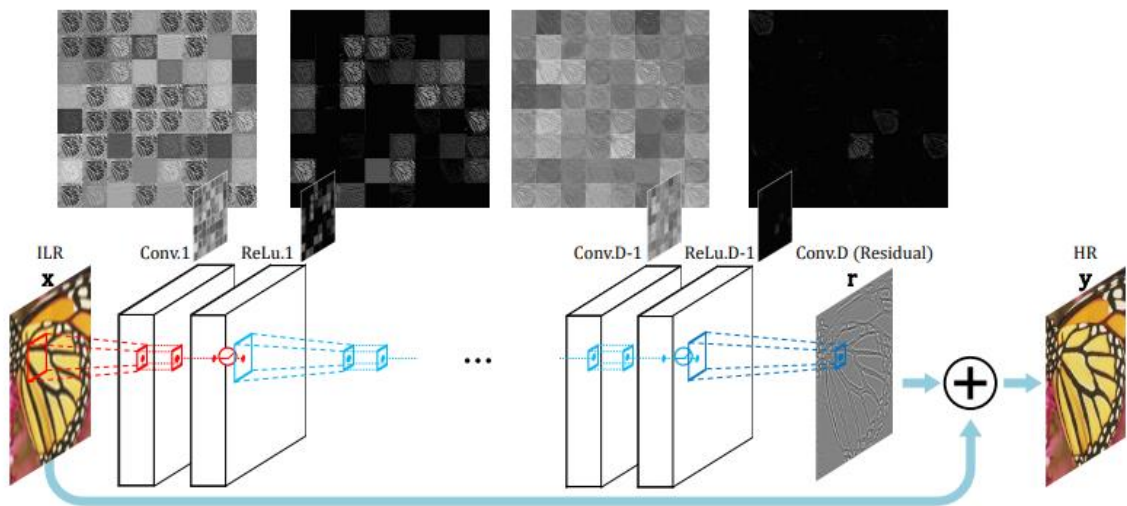


그림 3. VDSR의 모델 구조[3]

나. FSRCNN

SRCNN의 저자인 Dong은 SRCNN을 개선한 FSRCNN[2]을 제안하였다. 기존 SRCNN의 경우, 입력 LR 영상을 HR 영상의 크기로 바이큐빅 알고리즘을 이용하여 보간한 뒤, 이를 CNN에 입력하였다. 이로 인해 매우 CNN을 통한 연산 복잡도가 증가하게 되었다. Dong은 이 문제를 개선하기 위하여 입력영상으로 LR 영상을 그대로 이용하는 대신, 특징 맵의 크기를 키우는 레이어는 Deconvolution 레이어를 이용하였다. 해당 레이어는 특징 맵의 크기를 증가시켜주는 역할을 하므로, 원하는 스케일링 팩터에 해당하도록 출력의 크기를 조절하는

것이 가능했다. 이를 통하여 Dong은 SRCNN 대비 적은 파라미터 개수와 더 높은 속도를 가지는 FSRCNN을 제안하게 되었다.

다. VDSR

VDSR[3]은 SRCNN과 FSRCNN과는 다르게, 잔차학습에 기반하여 설계된 모델이다. VDSR 발표 이전 레이어 개수가 많은, 즉 깊은 딥러닝 모델에서 발생하는 기울기 소실 문제를 해결하기 위해 제안된 잔차 학습을 이용하는 ResNet[4]이 Kaiming He에 의해 제안되었다. 이는 이전 모델들 대비 더 많은 레이어들을 가진 깊은 모델들에 대해서 파라미터 학습이 상대적으로 잘 되도록 하였으며, 이를 통해 레이어의 개수와 성능이 비례하게 되었다.

Kim 에 의해 제안된 VDSR 은 ResNet 에 기반하여 설계된 모델로, SRCNN 과 유사하게 사전에 기존 업샘플링 알고리즘을 이용하여 LR 영상을 HR 영상 크기로 보간한 뒤, 이를 입력으로 이용한다. 하지만 해당 영상을 그대로 CNN 의 입력으로 이용하여 출력되는 HR 영상을 획득하는 SRCNN 과 달리, VDSR 은 업샘플링된 LR 영상과 HR 영상의 잔차에 대해서 학습하게 된다. 따라서 실제 CNN 이 출력하는 것은 LR 영상과 HR 영상의 잔차이고, 여기에 다시 업샘플링된 LR 영상을 더해줌으로써 최종적으로 복원된 HR 영상을 획득하게 된다.

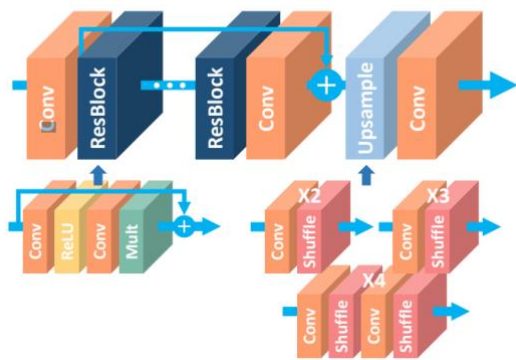


그림 4. EDSR의 모델 구조[5]

라. EDSR

EDSR[5]역시 ResNet 에 기반하여 설계된 모델로, Lim 에 의해 제안되었다. 해당 네트워크는 깊은 딥러닝 모델을 상대적으로 잘 학습시킬 수 있다는 잔차학습의 장점을 살려,

30 여개 이상의 레이어와, 기존 대비 4 배 이상의 레이어 채널수를 가지도록 설계되었다.

기존 모델들 대비 깊어지고 학습해야하는 파라미터가 늘어나면서, 저자는 학습의 안정성을 높이기 위하여 모델 내 레이어들을 여러 개의 블록으로 나눈 뒤, 각 블록의 결과물을 최종적으로 합쳐서 다음 레이어에서 이용하는 스킵 커넥션을 이용하였다.

또한 사전 업샘플링 방법 기반이었던 VDSR 과 달리 사후 업샘플링 방법에 기반하여, 네트워크 내에서 SPCNN(Sub Pixel Convolutional Neural Network)를 통해 입력된 LR 영상의 해상도를 HR 영상의 해상도로 업샘플링하였다. 이를 통해 VDSR 대비 효율적인 연산과정을 가지게 되었다.

마. SRGAN

SRGAN[6]은 GAN[7]에 기반하여 설계된 초해상화 모델로, Ledig 에 의해 제안되었다. 해당 네트워크는 이미지를 복원하는 생성 네트워크와, 학습 과정에서 HR 영상과 출력된 HR 영상을 구분하기 위한 분류 네트워크로 구성되어있다.

대다수의 초해상화 모델들이 학습 과정에서 MSE Loss 를 사용해왔으나, 이는 PSNR(Peak Signal-to-Noise Ratio)는 높지만, 영상의 형태가 달라 선명하게 보이지 않는 경우가 있었다. SRGAN 은 이를 해결하기 위하여 Perceptual loss 를 이용하여 PSNR 을 일부 포기하더라도 보다 선명한 영상을 출력하도록 학습되었다.

표 1. 모델들의 성능 비교표

Dataset	Scale	Bicubic PSNR/SSIM	SRCNN PSNR/SSIM	FSRCNN PSNR/SSIM	VDSR PSNR/SSIM	EDSR PSNR/SSIM	SRGAN PSNR/SSIM
Set5	x2	33.66/0.9299	36.66/0.9542	37.00/-	37.53/0.9587	38.20/0.9606	-
	x3	30.39/0.8682	32.75/0.9090	33.16/-	33.66/0.9213	34.76/0.9290	-
	x4	28.42/0.8104	30.48/0.8628	30.71/-	31.35/0.8838	32.62/0.8984	29.40/0.8472
Set14	x2	30.24/0.8688	32.42/0.9063	32.63/-	33.03/0.9124	34.02/0.9204	-
	x3	27.55/0.7742	29.28/0.8209	29.43/-	29.77/0.8314	30.66/0.8481	-
	x4	26.00/0.7027	27.49/0.7503	27.59/-	28.01/0.7674	28.94/0.7901	26.02/0.7397
BSD100	x2	29.56/0.8431	31.36/0.8879	31.80/-	31.90/0.8960	32.37/0.9018	-
	x3	27.21/0.7385	28.41/0.7863	28.60/-	28.82/0.7976	29.32/0.8104	-
	x4	24.46/0.6675	26.90/0.7101	26.98/-	27.29/0.7251	27.79/0.7437	25.16/0.6688

III. SUPER-RESOLUTION 기술 성능 비교

이번 섹션에서는 딥러닝 기반 초해상화 모델들의 성능을 비교 평가한다. 본 평가는 딥러닝 기반 초해상화 모델들의 벤치마크에 주로 이용되는 데이터셋인 Set5, Set14, BSD100에 대하여 스케일링 팩터 x2, x3, x4로 출력된 HR 영상 결과를 비교한다.

본 성능 비교 평가에 이용되는 메트릭으로는 화질의 손실정보를 나타내는 지표인 PSNR과 두 영상간의 유사도를 나타내는 지표인 SSIM을 사용하였다.

표 1을 통해 각 모델들의 PSNR 및 SSIM 결과를 확인할 수 있다. FSRCNN의 경우 SSIM 결과가, SRGAN의 경우 스케일링 팩터 x2, x3 결과가 제공되지 않아 제외하였다.

표 1을 통해 EDSR 모델이 모든 데이터셋의 모든 스케일링 팩터에 대해 가장 좋은 PSNR 및 SSIM 성능을 보여주고 있다는 것을 확인할 수 있다.

SRGAN의 경우 PSNR 및 SSIM 성능이 SRCNN과 비교하여도 크게 높지 않지만, MOS 평가 결과 상기 모델들 중 가장 높은 점수를 보였다.

IV. 결론

본 보고서에서는 심층 신경망을 이용한 초해상화 알고리즘의 동향에 대해서 분석하였다. 초해상화 방식에 따라 3가지로 분류하여 각 방식에 대하여 설명하였으며, 해당 방식들에 속하는 모델 5개에 대해서 평가하였다.

심층 신경망 기반 초해상화 기법은 전통적인 알고리즘 대비 좋은 결과를 보여주고 있지만, 높은 파라미터 수 및 연산량으로 인하여 사실상 GPU의 이용이 강제되고 있는 상황이다. 향

후 최적화를 통해 상용화될 수 있는 좋은 성능의 모델이 연구되는 것을 기대할 수 있다.

참고문헌

- [1] Chao Dong, Chen Change Loy, Kaiming He, Xiaoou Tang. Image Super-Resolution Using Deep Convolutional Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), Preprint, 2015
- [2] Chao Dong, Chen Change Loy, Xiaoou Tang. Accelerating the Super-Resolution Convolutional Neural Network, in Proceedings of European Conference on Computer Vision (ECCV), 2016
- [3] Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks", Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 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] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [6] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [7] Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems 27 (2014).