

Super-Resolution 최신 기술 동향

Trends on Super-Resolution

지능기전공학부

최용호(21110389)

Abstract

Super-Resolution 는 딥러닝 기술이 발전하기 이전부터 다양한 필터들과 보간법을 통해서 영상의 고주파 성분, 저주파 성분 그리고 가우시안, 영상 압축 에러와 같은 다양한 열화들을 원본 영상으로 복원시키는 하나의 분야로써 연구되었다. 하지만, 폭발적인 데이터의 증가와 GPU의 발전은 멈춰있던 인공지능의 바람을 일으켰고, 현재 다양한 분야들과 함께 Super-Resolution의 분야는 Convolution neural networks(CNNs) 기반으로 하여 많은 연구가 활발히 진행중이다. 최근에는 GAN(Generative Adversarial Network), Transformer 등과 같은 다른 분야의 네트워크 구조나 새로운 손실 함수를 활용하여 더 좋은 성능을 내기도 하였다. 위 보고서에는 기본적인 Super-Resolution 의 구조에 대해서 알아보고 기존 CNN에서 Super-Resolution에 좋은 성능을 내는 효율적으로 레이어를 쌓는 기술과 이에 기반한 다양한 CNN모델들, 마지막으로 SOTA(state-of-art)에 해당하는 여러 논문 에 대한 동향을 파악하고자 한다.

I. Introduction

Super-Resolution은 영상 복원(Image Restoration)의 한 분야로서 작은 해상도(Low Resolution : LR)가 가진 제한된 정보를 가지고 얼마나 원본(High Resolution : HR)에 근접한 영상을 복원할 수 있는지에 대한 문제를 가지고있다. 이 분야의 활용으로는 Fig.1과 같은 가우시안 에러, 소금 후추 에러 그리고 광원에 의한 빛번짐 등과 같은 영상을 취득할 때에 생길 수 있는 열화와 영상을 전송하거나 매체에 저장할 때 일어나는 영상 압축 열화들(Compression Artifacts)에 대해서도 위 분야를 이용하여 real-world(medical images [1], [2], [3], surveillance and security [4],[5])에 대한 다양한 연구 또한 이루어지고 있다.

위와 같은 실질적인 상용화를 위해서 Super-Resolution 분야에서는 원본 영상을 Down-Sampling한 저해상도(Low Resolution) 영상에 다양한 열화를 부여하여 다시 원본 영상(High Resolution)으로 복원시키는 Fig.2 [6] 와 같은 프레임구조를 가지고있다.

이전의 Super-Resolution은 딥러닝 발전 이전부터 원론적인 방법으로는 연구가 많이 되어있었다. 이웃한 픽셀들을 활용한 보간법(cubic, Lanczos Interpolation)[7],[8] 통계적인 방법 [9] 그리고 에지 검출 기법[10] 등등과 같은 방법들로 많은 원론적인 방법이 연구가 되었었다. 하지만 CNN(Convolution Neural Network)의 커널 연산을 통해 영상들의 특징을 알아내어 물체 검출, 물체 분류 및 물체 추적 심지어 영상 생성과 같은 다양한 영상 처리 분야에 활용되고 있다. 이에 따른 SRCNN과 FSRCNN [11],[12]모델은 이러한 CNN을 통해서 Super-Resolution 구조의 근본이 되는 프레임워크를 만들어냈고, 현재까지도 베이스 구조를 따른 다양한 모델 연구를 하고있다.

초반 Super resolution의 CNN 계층은 영상을 Up-sampling시킬 수가 없어서 SRCNN의 경우 원본의 해당하는 영상을 Bicubic interpolation의 원론적인 방법의 Up-sampling을 통하여 LR을 HR의 크기로 복원을 시킨다음 L1_loss를 통한 Y에 해당하는 HR과의 MSE를 줄여나가면서 학습시킨다. 이에 더하여 FSRCNN 에서는 이러한 한계를 Deconvolution(Transpose Convolution) 이라는 레이어로 Up-sampling이 가능하게 되면서 LR을 그대로 사용할 수 있어서 오직 CNN 네트워크만을 활용한



Fig.1 Image Artifacts

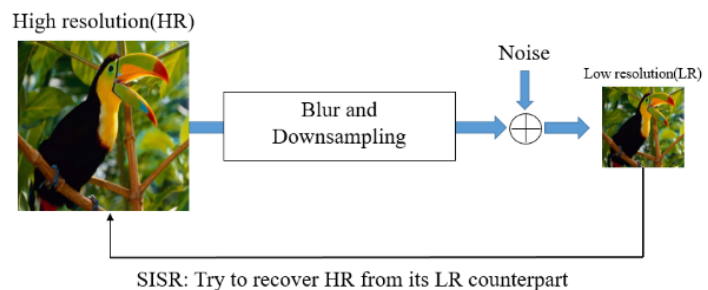


Fig.2 SuperResolution Base Structure

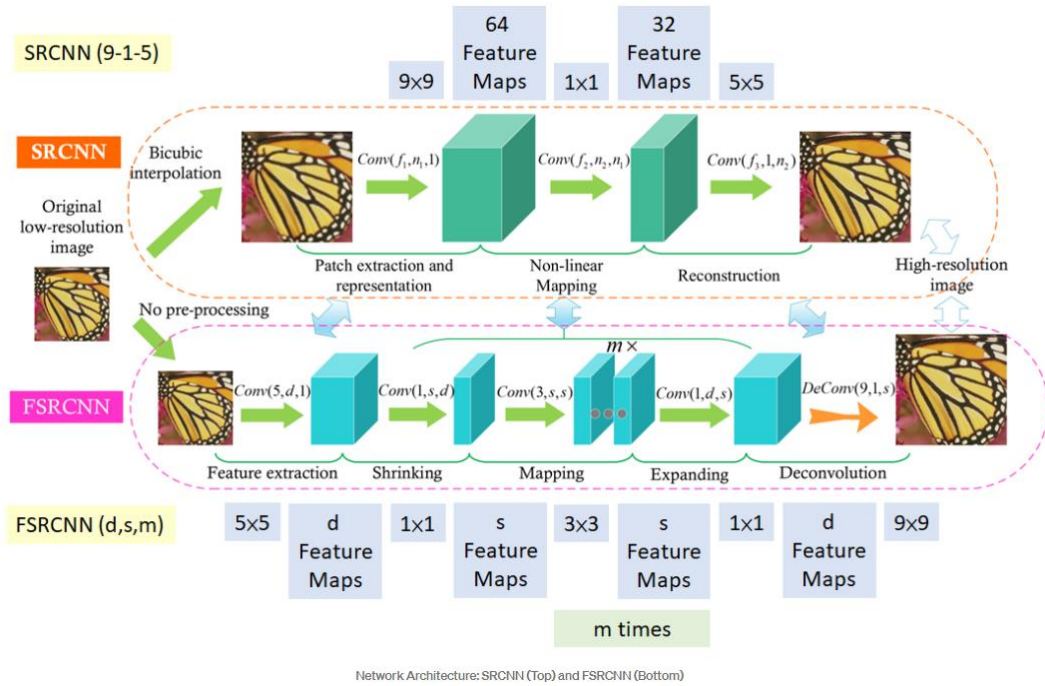


Fig.3 SRCNN, FSRCNN Structure

End-to-End 모델들이 나오기 시작되었다. 그러나, Super-Resolution에는 이러한 Deconvolution에 중복된 연산에 의한 Checkboard Artifact[13]가 존재한다. 이는 Deconvolution의 파라미터에 의해 완화될 수 있지만, 너무 많은 실험을 요구한다. 이에 대하여 LR 특징맵을 영상에 맞게 재배열하는 Sub-pixel Shuffle layer[14]가 나왔으며 완전하게 Checkboard Artifact를 제거하지는 못하지만 Deconvolution 보다는 복잡도와 성능에서 좋은 평을 받고있어서 Deconvolution보다는 많이 사용되는 Up-sampling Layer로 이용되고있다. SRCNN과 FSRCNN 모델은 단순히 영상의 특징맵을 추출하는데, 이는 이전 영상의 특징맵을 손실한다는 단점이 존재하였고, 이는 CNN 분류 모델 ResNet[15]에서 착안된 Low-level의 영상의 특징맵을 사용하거나, 중간 레이어의 특징맵을 활용하는 Skip Connection 이라 일컫는 Residual Learning, Channel attention 그리고 Dense connections (Fig.4) 등과 같은 기술들로 Super-Resolution모델의 다양한 변화와 개선을 이끌어내었다.

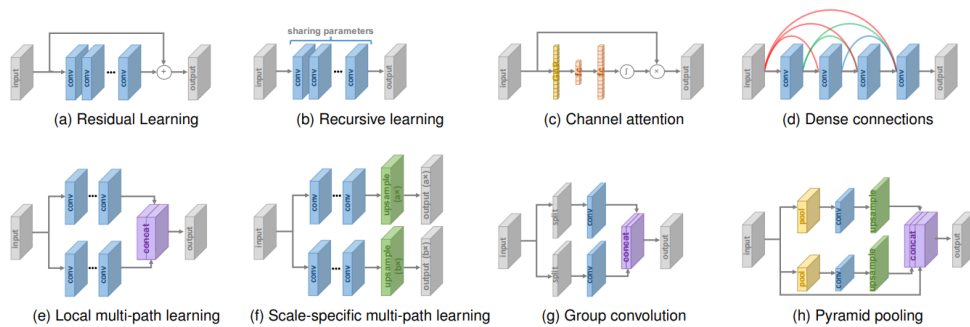


Fig.4 Efficient Network design strategies in Super-Resolution [16]

최근에는 이안 굿 펠로우 저자가 제안했던 GAN(Generative Adversarial Network)[17]로 CNN모델로 구성된 생성자(Generator), 구별자(Discriminator)를 이용한 적대적인 학습을 통해서 영상 도메인을 변경하는 분야(Image Translation)를 활용하여 Super-Resolution 분야에도 활용되고있으며, 더 나아가 자연어 처리(Natural Language Processing)에서 Attention 기법만을 이용한 Transformer[18] 는 영상 분야에도 적용시키면서 물체 검출에서도 활발하게 연구되고 있으며, 이어서 Super-Resolution 분야에서도 SOTA의 성능을 보여주고 있다.

따라서, 지금까지 기본적인 Super-Resolution 베이스 모델의 발전 과정을 알아보았고 2장에서는 CAR, Transformer와 같은 다른 변화를 가진 프레임구조의 SOTA 모델을 살펴본다. 이어서 3장에서는 Super-Resolution의 평가 지표에 대해 분석을 해보고, 결론에서는 이전 장을 기반으로 최신 모델의 어떠한 부분이 Super-Resolution의 개선을 이끌어냈는지 알아보도록 한다.

II. Trends on Super-Resolution

1. SwinIR : Super Resolution based Transformer

SwinIR은 Transformer 모델이 비전 영역으로 확대되면서 발전하였다. ViT(Vision Transformer)를 시작으로 고정된 패치 영상을 embedding을 통해서 Transformer를 적용하였지만, 고정된 패치라는 점의 한계를 Swin Transformer (Shifted Windows) 고정된 패치를 움직이면서 영상 내의 연관성을 더 잘 학습시킬 수 있는 장점을 가지고 영상 복원 분야에 적용한 경우이다.

위 Fig.5 모델은 일반적인 Super-Resolution의 지도학습의 형태를 띄고있으며 L1 손실함수를 이용하여 학습시킨다. Shallow Feature Extraction 에서는 일반적인 Convolution 하나의 레이어를 이용하며, 이 후에 low-level 을 connection을 통해 Deep Feature Extraction 이 후에 합쳐준다. Deep Feature Extraction 에서는 다수의 Swin Transformer 와 Convolution 으로 되어있는 다수의 RSTB(Residual Swin Transformer Block) 로 이루어져있고 Swin Transformer 과 Convolution의 사이에는 각각 Transformer로 피쳐들을 보내주기위한 embedding과 unembedding이 존재한다. 마지막단에는 일반적인 Sub-pixel Up-sampling 레이어를 이용하고 차원을 영상에 맞춰주기위한 하나의 convolution 레이어로 이루어져 있다.

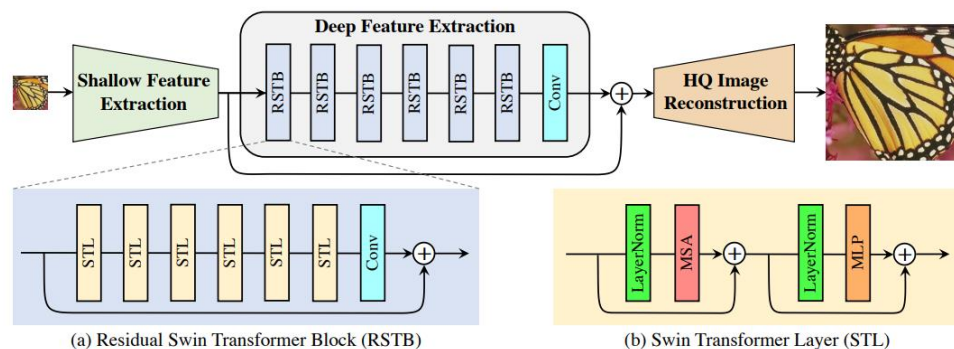


Fig.5 The architecture of the proposed SwinIR for image restoration.

성능면에서는 현재 영상 복원 분야에서 SOTA를 기록하고 있으며, 모델의 복잡도가 크지 않다는 점에서 흥미롭다. 이는 Convolution 의 레이어 보다 Transformer 의 구조가 영상 복원에 어쩌면 더 적합하다는 의미가 되는데 학습 시간에 대해서는 Shifted Windows 의 알고리즘 복잡도가 높기 때문에 오래 걸린다는 단점이 있다.

Method	Scale	Training Dataset	Set5 [3]		Set14 [87]		BSD100 [58]		Urban100 [34]		Manga109 [60]	
			PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RCAN [95]	×2	DIV2K	38.27	0.9614	34.12	0.9216	32.41	0.9027	33.34	0.9384	39.44	0.9786
SAN [15]	×2	DIV2K	38.31	0.9620	34.07	0.9213	32.42	0.9028	33.10	0.9370	39.32	0.9792
IGNN [100]	×2	DIV2K	38.24	0.9613	34.07	0.9217	32.41	0.9025	33.23	0.9383	39.35	0.9786
HAN [63]	×2	DIV2K	38.27	0.9614	34.16	0.9217	32.41	0.9027	33.35	0.9385	39.46	0.9785
NLSA [61]	×2	DIV2K	38.34	0.9618	34.08	0.9231	32.43	0.9027	33.42	0.9394	39.59	0.9789
SwinIR (Ours)	×2	DIV2K	38.35	0.9620	34.14	0.9227	32.44	0.9030	33.40	0.9393	39.60	0.9792
SwinIR+ (Ours)	×2	DIV2K	38.38	0.9621	34.24	0.9233	32.47	0.9032	33.51	0.9401	39.70	0.9794
DBPN [31]	×2	DIV2K+Flickr2K	38.09	0.9600	33.85	0.9190	32.27	0.9000	32.55	0.9324	38.89	0.9775
IPT [9]	×2	ImageNet	38.37	-	34.43	-	32.48	-	33.76	-	-	-
SwinIR (Ours)	×2	DIV2K+Flickr2K	38.42	0.9623	34.46	0.9250	32.53	0.9041	33.81	0.9427	39.92	0.9797
SwinIR+ (Ours)	×2	DIV2K+Flickr2K	38.46	0.9624	34.61	0.9260	32.55	0.9043	33.95	0.9433	40.02	0.9800
RCAN [95]	×3	DIV2K	34.74	0.9299	30.65	0.8482	29.32	0.8111	29.09	0.8702	34.44	0.9499
SAN [15]	×3	DIV2K	34.75	0.9300	30.59	0.8476	29.33	0.8112	28.93	0.8671	34.30	0.9494
IGNN [100]	×3	DIV2K	34.72	0.9298	30.66	0.8484	29.31	0.8105	29.03	0.8696	34.39	0.9496
HAN [63]	×3	DIV2K	34.75	0.9299	30.67	0.8483	29.32	0.8110	29.10	0.8705	34.48	0.9500
NLSA [61]	×3	DIV2K	34.85	0.9306	30.70	0.8485	29.34	0.8117	29.25	0.8726	34.57	0.9508
SwinIR (Ours)	×3	DIV2K	34.89	0.9312	30.77	0.8503	29.37	0.8124	29.29	0.8744	34.74	0.9518
SwinIR+ (Ours)	×3	DIV2K	34.95	0.9316	30.83	0.8511	29.41	0.8130	29.42	0.8761	34.92	0.9526
IPT [9]	×3	ImageNet	34.81	-	30.85	-	29.38	-	29.49	-	-	-
SwinIR (Ours)	×3	DIV2K+Flickr2K	34.97	0.9318	30.93	0.8534	29.46	0.8145	29.75	0.8826	35.12	0.9537
SwinIR+ (Ours)	×3	DIV2K+Flickr2K	35.04	0.9322	31.00	0.8542	29.49	0.8150	29.90	0.8841	35.28	0.9543
RCAN [95]	×4	DIV2K	32.63	0.9002	28.87	0.7889	27.77	0.7436	26.82	0.8087	31.22	0.9173
SAN [15]	×4	DIV2K	32.64	0.9003	28.92	0.7888	27.78	0.7436	26.79	0.8068	31.18	0.9169
IGNN [100]	×4	DIV2K	32.57	0.8998	28.85	0.7891	27.77	0.7434	26.84	0.8090	31.28	0.9182
HAN [63]	×4	DIV2K	32.64	0.9002	28.90	0.7890	27.80	0.7442	26.85	0.8094	31.42	0.9177
NLSA [61]	×4	DIV2K	32.59	0.9000	28.87	0.7891	27.78	0.7444	26.96	0.8109	31.27	0.9184
SwinIR (Ours)	×4	DIV2K	32.72	0.9021	28.94	0.7914	27.83	0.7459	27.07	0.8164	31.67	0.9226
SwinIR+ (Ours)	×4	DIV2K	32.81	0.9029	29.02	0.7928	27.87	0.7466	27.21	0.8187	31.88	0.9423
DBPN [31]	×4	DIV2K+Flickr2K	32.47	0.8980	28.82	0.7860	27.72	0.7400	26.38	0.7946	30.91	0.9137
IPT [9]	×4	ImageNet	32.64	-	29.01	-	27.82	-	27.26	-	-	-
RRDB [81]	×4	DIV2K+Flickr2K	32.73	0.9011	28.99	0.7917	27.85	0.7455	27.03	0.8153	31.66	0.9196
SwinIR (Ours)	×4	DIV2K+Flickr2K	32.92	0.9044	29.09	0.7950	27.92	0.7489	27.45	0.8254	32.03	0.9260
SwinIR+ (Ours)	×4	DIV2K+Flickr2K	32.93	0.9043	29.15	0.7958	27.95	0.7494	27.56	0.8273	32.22	0.9273

Fig.6 Quantitative comparison (average PSNR/SSIM) for classical image SR

2. CAR : Super Resolution based Down-Sampling Resampler Network

지금까지는 Fig.2 와 같은 원론적인 방법론으로 Down-Sampling을 진행하고 이를 딥러닝 모델을 이용하여 복원을 시켰다면, 위 논문에서는 원론적인 Down-Sampling의 방법을 주로 가장 가까운 픽셀의 값을 활용한 보간법중에서도 bicubic을 주로 사용하는데 이는 한 픽셀을 생성하는데 4개의

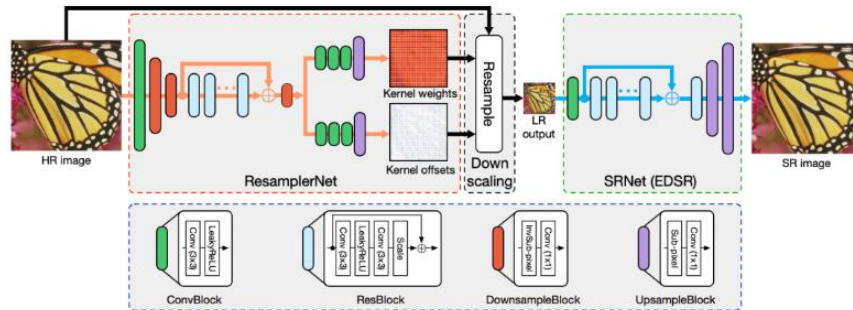


Fig.7 The architecture of the proposed CAR Model

값을 사용하여 어떠한 영상에서는 영상의 에지가 훼손되며 정확한 예측이 불가하다는 것과 이전의 Down-sampling CNN네트워크의 기하학적인 패턴의 한계를 근거로 Deformable CNN의 아이디어를 활용한 딥러닝 기반의 적응적인 가중치 필터로 LR생성하는 Down-Sampling 방법론을 제시하였다. 다음 Fig.7은 모델의 구조를 보여준다. InvSubPixel 레이어들과 Residual block 이전의 피쳐맵을 connection을 통해 HR의 유의미한 특징을 잡고 적응적인 가중치 필터를 생성하기위해 나뉜진 두 개의 SubPixel Up-sampling 레이어를 통해 스케일에 의해 HR 크기의 피쳐맵 커널을 생성한다. 이 커널은 수평과 수직의 정해진 수식에 의해 픽셀값이 예측이 되어 Resampling 된다.

또한, 위 모델은 비지도학습으로 딥러닝 기반의 커널로 생성된 LR을 다른 Super-Resolution Network에 적용하여 나온 예측된 HR 을 L1 손실함수와 원본적인 방법과 동일하게 resampled 커널의 결과에는 중심에 몰려있는 경우가 있어서 이를 막기 위한 offset loss 와 TV(Total Variation) loss 를 사용하여 학습을 진행한다. 결과(Table,[])는 Bicubic 에 비해 약 1db이상의 성능이 올라가는 놀라운 성능을 볼 수 있다. 위 모델은 새로운 전략의 Super-Resolution의 구조를 생성하였고, 점차 발전해나가는 Super-Resolution의 다양한 모델에 맞는 Down-sampling 커널을 생성한다는 장점이 있다.

Table 1: Quantitative evaluation results (PSNR / SSIM) of different image downscaling methods for SR on benchmark datasets: Set5, Set14, BSD100, Urban100 and DIV2K (validation set).

Upscaling		Bicubic		EDSR				
Downscaling		Bicubic	CAR	Bicubic	Perceptually	DPID	L0-regularized	CAR
Set5	2x	33.66 / 0.9299	34.38 / 0.9426	38.06 / 0.9615	31.45 / 0.9212	37.02 / 0.9573	35.40 / 0.9514	38.94 / 0.9658
	4x	28.42 / 0.8104	28.93 / 0.8335	32.35 / 0.8981	25.65 / 0.7805	31.75 / 0.8913	31.05 / 0.8847	33.88 / 0.9174
Set14	2x	30.24 / 0.8688	31.01 / 0.8908	33.88 / 0.9202	29.26 / 0.8632	32.82 / 0.9119	31.56 / 0.9008	35.61 / 0.9404
	4x	26.00 / 0.7027	26.39 / 0.7326	28.64 / 0.7885	24.21 / 0.6684	28.27 / 0.7784	27.67 / 0.7702	30.31 / 0.8382
B100	2x	29.56 / 0.8431	30.18 / 0.8714	32.31 / 0.9021	28.62 / 0.8383	31.47 / 0.8922	30.75 / 0.8816	33.83 / 0.9262
	4x	25.96 / 0.6675	26.17 / 0.6963	27.71 / 0.7432	24.61 / 0.6391	27.27 / 0.7341	27.00 / 0.7293	29.15 / 0.8001
Urban100	2x	26.88 / 0.8403	27.38 / 0.8620	32.92 / 0.9359	26.39 / 0.8483	31.64 / 0.9271	30.23 / 0.9172	35.24 / 0.9572
	4x	23.14 / 0.6577	23.35 / 0.6844	26.62 / 0.8041	21.58 / 0.6295	26.07 / 0.7967	25.83 / 0.7957	29.28 / 0.8711
DIV2K (validation)	2x	31.01 / 0.9393	33.18 / 0.9317	36.76 / 0.9482	31.23 / 0.8984	35.75 / 0.9419	34.69 / 0.9354	38.26 / 0.9599
	4x	26.66 / 0.8521	28.50 / 0.8557	31.04 / 0.8452	26.28 / 0.7381	30.53 / 0.8373	30.18 / 0.8340	32.82 / 0.8837

Note: Red color indicates the best performance and Blue color represents the second.

Table 2: Evaluation results (PSNR / SSIM) of 4× upscaling using different SR networks on benchmark images downsampled by the CAR model.

Upscaling	Downscaling	Set5	Set14	B100	Urban100	DIV2K
SRDenseNet	Bicubic	32.02 / 0.8934	28.50 / 0.7782	27.53 / 0.7337	26.05 / 0.7819	- / -
	CAR†	33.16 / 0.9067	29.85 / 0.8201	28.73 / 0.7794	27.97 / 0.8403	32.24 / 0.8674
	CAR‡	32.63 / 0.9047	29.24 / 0.8122	28.44 / 0.7781	27.12 / 0.8248	31.77 / 0.8654
D-DBPN	Bicubic	32.47 / 0.8980	28.82 / 0.7860	27.72 / 0.7400	26.38 / 0.7946	- / -
	CAR†	33.07 / 0.9061	29.75 / 0.8189	28.70 / 0.7789	27.98 / 0.8376	32.13 / 0.8664
	CAR‡	32.71 / 0.9055	29.17 / 0.8076	28.45 / 0.7784	27.00 / 0.8222	31.76 / 0.8650
RDN	Bicubic	32.47 / 0.8990	28.81 / 0.7871	27.72 / 0.7419	26.61 / 0.8028	- / -
	CAR†	33.34 / 0.9132	29.93 / 0.8308	28.89 / 0.7961	28.53 / 0.8582	32.32 / 0.8756
	CAR‡	33.15 / 0.9112	29.59 / 0.8227	28.79 / 0.7913	27.69 / 0.8412	32.20 / 0.8747
RCAN	Bicubic	32.63 / 0.9002	28.87 / 0.7889	27.77 / 0.7436	26.82 / 0.8087	30.77 / 0.8459
	CAR†	33.84 / 0.9187	30.27 / 0.8383	29.16 / 0.8021	29.23 / 0.8719	32.81 / 0.8842
	CAR‡	33.37 / 0.9138	29.87 / 0.8294	28.95 / 0.7953	28.28 / 0.8541	32.46 / 0.8786

CAR†: the CAR model is trained jointly with its corresponding SR model.

CAR‡: the SR model is trained using the downsampled images generated by the CAR model that is jointly with the EDSR.

Note: Red color indicates the best performance and Blue color represents the second. The '-' indicates that results are not provided by the corresponding original publication.

Table.1 Result of CAR

III. Super-Resolution Evaluation Score

영상 화질 개선(Image Restoration) 분야에서는 PSNR(Peak Signal-to-Noise Ratio) 신호 잡음 대비를 전통적인 화질 평가지표로 사용하고 있으며, 이는 수식을 살펴보면, 픽셀의 최대값에 대비하여 원본 영상과 복원 영상 픽셀 간의 오차 값(MSE)을 사용하여 복원 영상이 얼마나 에러 값을 픽셀 단위에서 확인할 수 있는 객관적인 지표이다. 하지만, 오늘날에는 Fig.8과 같이 PSNR이 영상 평가 지표의 전부는 아니라는 평가가 논의되어, 주관적인 평가와 함께 밝기(Luminance), 대조(Contrast) 그리고 밝기와 대조의 관계를 공분산으로 나타낸 구조(Structure)인 3가지 측면에서 영상을 평가하는 SSIM (Structural similarity index) 또한 중요한 지표로 사용되고 있다. Fig.9

$$\begin{aligned}
 PSNR &= 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) & MSE &= \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \\
 &= 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \\
 &= 20 \log_{10}(MAX_I) - 10 \log_{10}(MSE)
 \end{aligned}$$

$$\begin{aligned}
 SSIM(A, B) &= l(A, B)c(A, B)s(A, B) \\
 &= \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)}
 \end{aligned}$$

$$\begin{aligned}
 l(A, B) &= \frac{2\mu_A\mu_B + C_1}{\mu_A^2 + \mu_B^2 + C_1} \\
 c(A, B) &= \frac{2\sigma_{AB} + C_2}{\sigma_A^2 + \sigma_B^2 + C_2} \\
 s(A, B) &= \frac{\sigma_{AB} + C_3}{\sigma_A\sigma_B + C_3}
 \end{aligned}$$

Fig.8 Formula of PSNR and SSIM

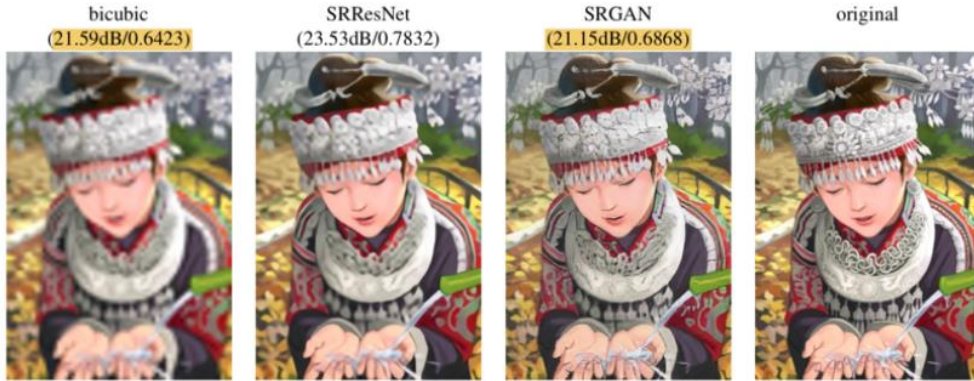


Fig.9 Example about relation between PSNR and SSIM

Super-Resolution에서 중요하게 생각할 부분은 객관적 화질 평가 지표 이 외에도 네트워크의 복잡도 또한 중요하게 생각한다. 레이어를 쌓으면 화질 개선이 일어난다는 것은 어떻게 보면 당연한 결과이고, 모델이 무거워질수록 real-time과는 trade-off 관계가 발생하기 때문에 이는 객관적 화질 평가 지표와 함께 모델을 평가하는데 다음 Fig.10 같이 적합한 평가 방법으로 활용되고 있다.

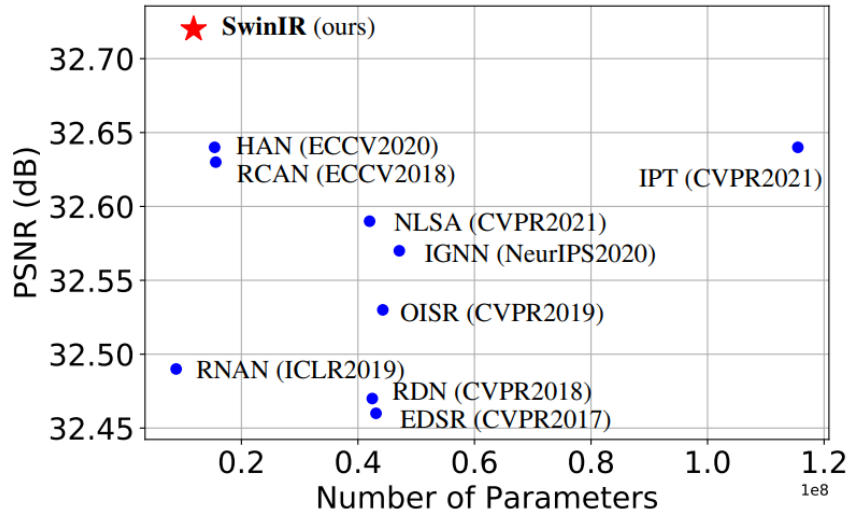


Fig.10 Graph of SOTA in SwinIR

IV. Conclusion

지금까지 Super-Resolution의 End-to-End 과정과 이를 베이스를 한 CNN모델 및 효율적인 기술들과 여러 분야를 활용한 Super-Resolution의 기술들에 대해서도 다뤄보았다. 여태 보았듯이 CNN의 레이어를 쌓는 부분에서 성능을 이끌어내는데 한계가 존재한다. 이에 따라 다양한 분야들이 생겨나고, 기존의 틀에 벗어난 새로운 모델이 나오면 서로의 분야에 적용해보는 연구가 활발하게 진행되고 있다. 따라서, 앞으로의 연구자들은 기본적인 베이스 모델에 국한되는 편협한 시각에 벗어나서, 다른 분야에서 활용된 모델에도 관심을 갖고 적용하려는 열린 사고를 갖는 것이 더욱 중요할 것이다.

V. Reference

- [1] H. Greenspan, "Super-resolution in medical imaging," The Computer Journal, vol. 52, 2008.
- [2] J. S. Isaac and R. Kulkarni, "Super resolution techniques for medical image processing," in ICTSD, 2015.
- [3] Y. Huang, L. Shao, and A. F. Frangi, "Simultaneous superresolution and cross-modality synthesis of 3d medical images using weakly-supervised joint convolutional sparse coding," in CVPR, 2017
- [4] L. Zhang, H. Zhang, H. Shen, and P. Li, "A super-resolution reconstruction algorithm for surveillance images," Elsevier Signal Processing, vol. 90, 2010.
- [5] P. Rasti, T. Uiboupin, S. Escalera, and G. Anbarjafari, "Convolutional neural network super resolution for face recognition in surveillance monitoring," in AMDO, 2016.
- [6] Yang W, Zhang X, Tian Y, Wang W, Xue JH, Liao Q. Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia. 2019 May 28;21(12):3106-21.
- [7] R. Keys, "Cubic convolution interpolation for digital image processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, 1981

- [8] C. E. Duchon, "Lanczos filtering in one and two dimensions," *Journal of Applied Meteorology*, vol. 18, 1979
- [9] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior," *TPAMI*, vol. 32, 2010.
- [10] G. Freedman and R. Fattal, "Image and video upscaling from local self-examples," *TOG*, vol. 30, 2011.
- [11] Dong, C., Loy, C.C., He, K. and Tang, X., 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), pp.295-307.
- [12] Dong, C., Loy, C.C. and Tang, X., 2016, October. Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391-407). Springer, Cham.
- [13] Checkboard artifacts : Odena, Augustus, Vincent Dumoulin, and Chris Olah. "Deconvolution and checkerboard artifacts." *Distill* 1.10 (2016): e3.
- [14] Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D. and Wang, Z., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1874-1883).
- [15] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [16] Wang, Z., Chen, J. and Hoi, S.C., 2020. Deep learning for image super-resolution: A survey. *IEEE transactions on pattern analysis and machine intelligence*.
- [17] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2020. Generative adversarial networks. *Communications of the ACM*, 63(11), pp.139-144.
- [18] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).
- [19] Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L. and Timofte, R., 2021. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1833-1844).
- [20] Sun, W. and Chen, Z., 2020. Learned image downscaling for upscaling using content adaptive resampler. *IEEE Transactions on Image Processing*, 29, p p.4027-4040.