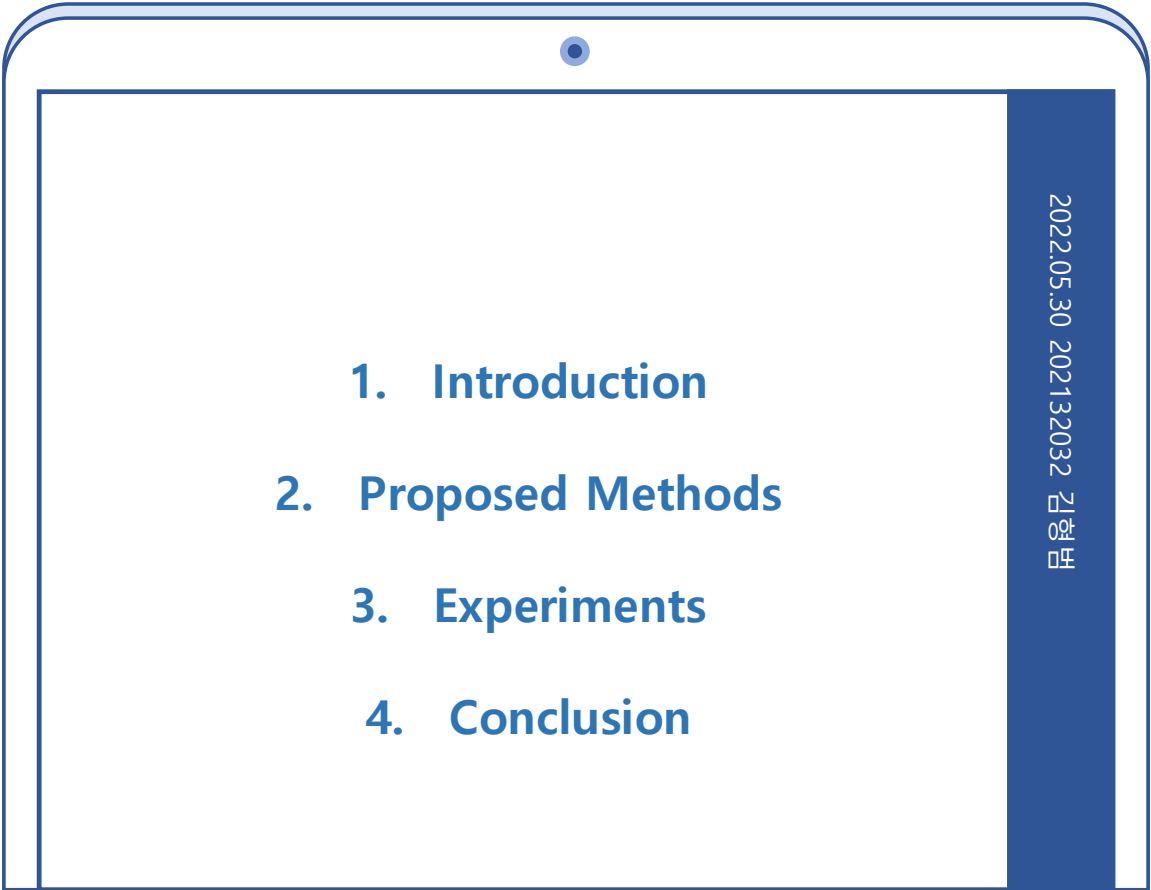


ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks

Wang, Xintao, et al. "EsrGAN: Enhanced super-resolution generative adversarial networks." *Proceedings of the European conference on computer vision (ECCV) workshops*. 2018.

2022.05.30 202132032 김희범

- 
1. Introduction
 2. Proposed Methods
 3. Experiments
 4. Conclusion

2022.05.30 202132032 김형범

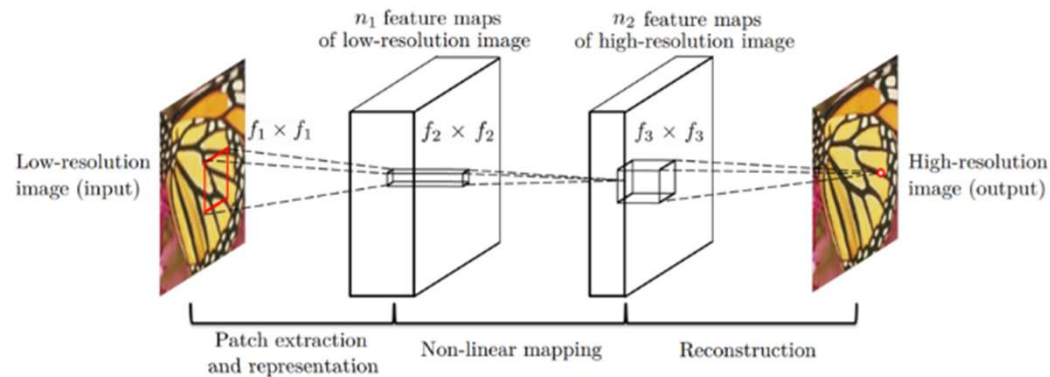
1. Introduction

2022.05.30 202132032 김희범

Introduction

○ SR task란?

- 단일 이미지 초해상도(SISR) 작업을 의미하며 단일 저해상도 이미지에서 고해상도 이미지를 복구하는 것을 목표로 한다.
- 연구 커뮤니티와 AI 기업들에서 점점 더 많은 관심을 끌고 있는 task 이다.
- 이미 많은 기술과 모델들이 나와있으며 성능은 주로 PSNR(최대 신호 대 잡음비)로 측정한다.



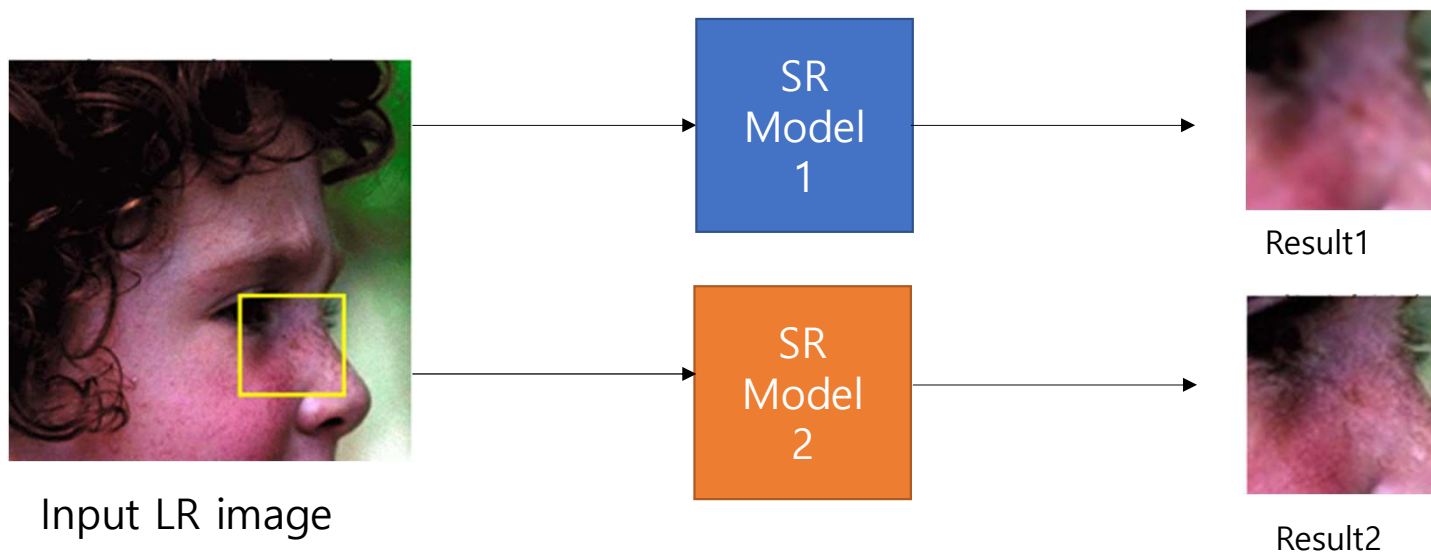
Introduction

◦ PSNR과 기존 모델들의 성능

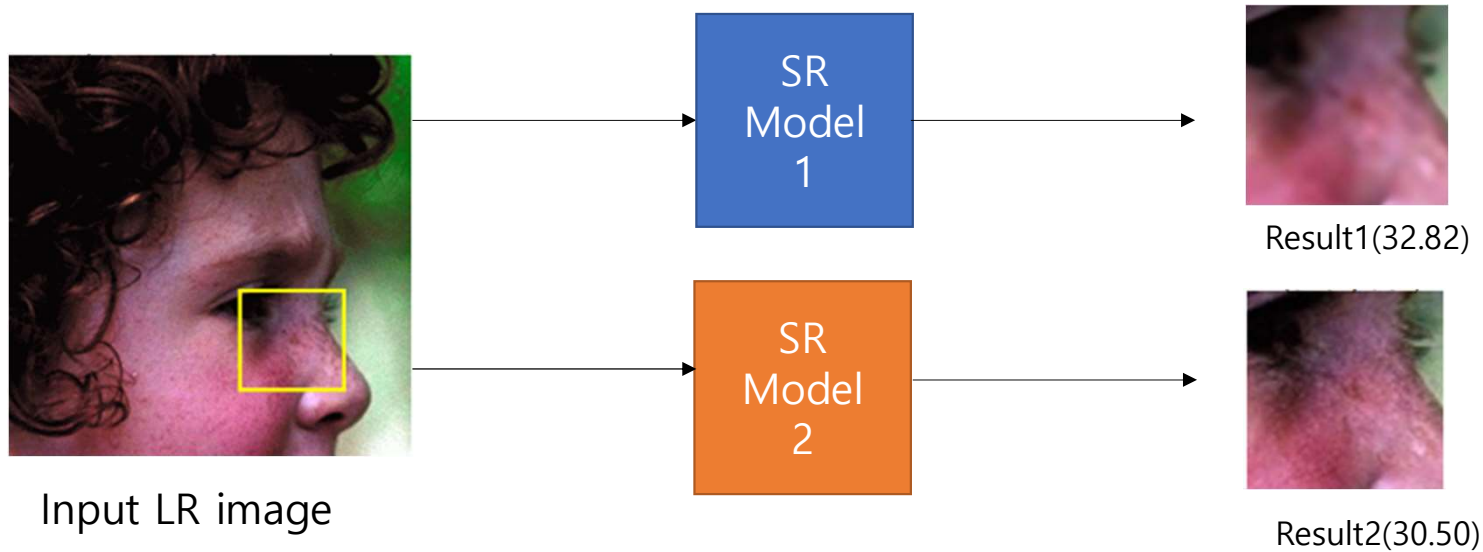
- PSNR(최대 신호 대 잡음비)는 신호가 가질 수 있는 최대 전력에 대한 잡음의 전력을 나타낸 것이다. 주로 영상 또는 동영상 손실 압축에서 화질 손실 정보를 평가할 때 사용된다.
- 기존의 많은 SR 모델(SRCNN, VDSR, EDSR, SRGAN 등)은 평가 척도를 PSNR로 설정하고 PSNR을 최대화하는 것을 목표로 두었고 높은 PSNR의 결과 이미지를 도출.

→ **PSNR 만으로 결과 이미지의 성능을 도출하는 것이 맞는 것일까?**

Introduction



Introduction



- PSNR 평가 기준이 인간의 지각적인 평가 기준과 다르기 때문에 PSNR이 높은 이미지가 더 안 좋게 보이는 경우가 있다.
- 고주파 디테일 없이 평활화된 결과 이미지들이 PSNR이 높게 측정되는 경향이 있다.

Introduction

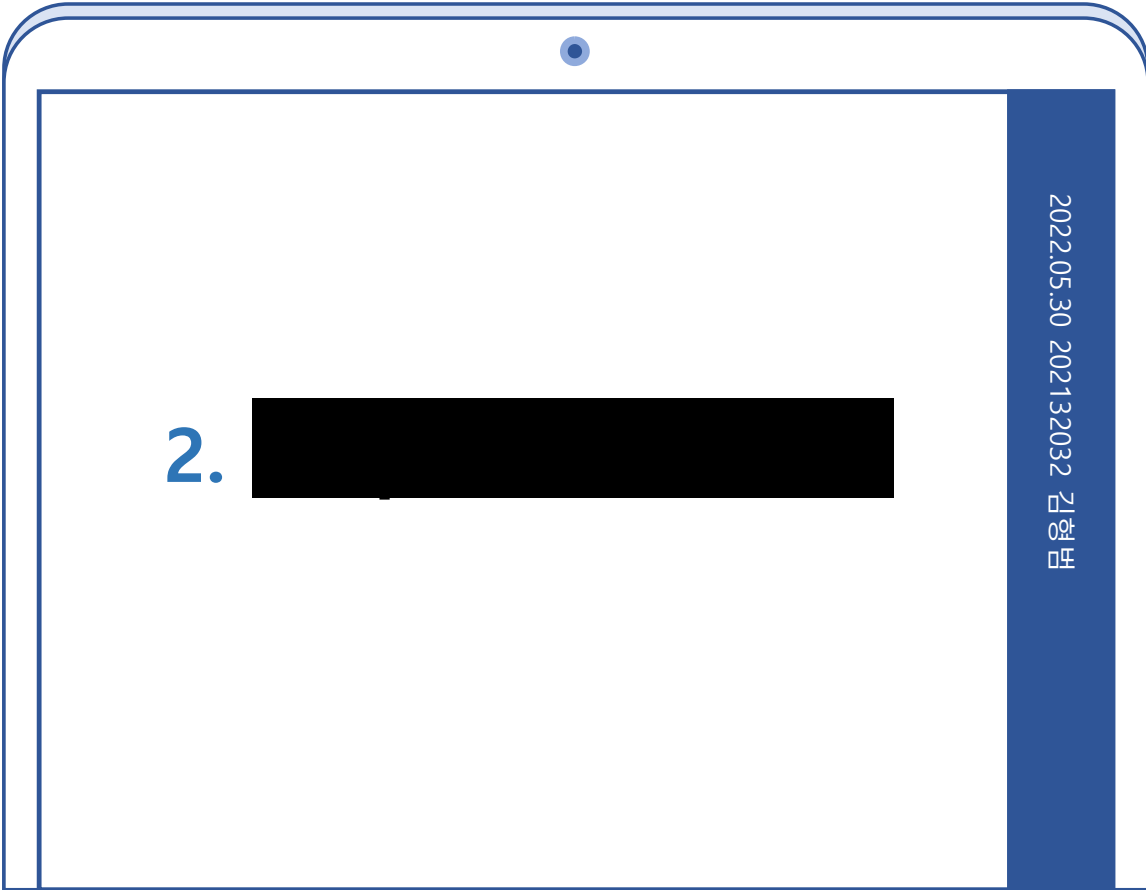
○ ESRGAN은 기존 기술 중 뛰어난 성능을 보이는 SRGAN을 개선하여 시각적으로 사실적이고 자연스러운 texture로 더 나은 시각적 품질을 달성했다.

- SRGAN의 구성요소를 살펴보고 세가지 측면에서 모델을 개선한다.

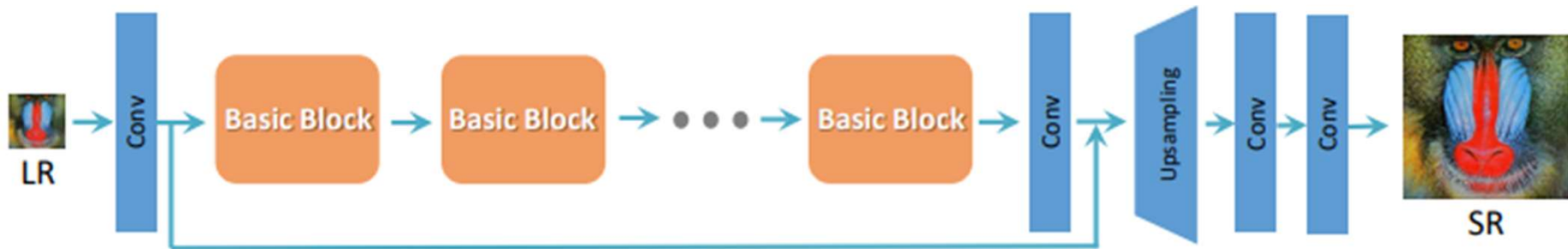
1. 용량이 더 크고 train에 용이한 RDDB(Residual-in-Residual Dense Block)를 도입하여 네트워크 구조를 개선한다.

2. batch normalization layer를 제거하고 Residual Scaling과 같은 더 작은 초기화 기법을 사용하여 매우 깊은 네트워크를 쉽게 훈련시킨다.

3. Activate 하기 전의 feature를 활용하는 개선된 perceptual loss를 제안한다.



Proposed method

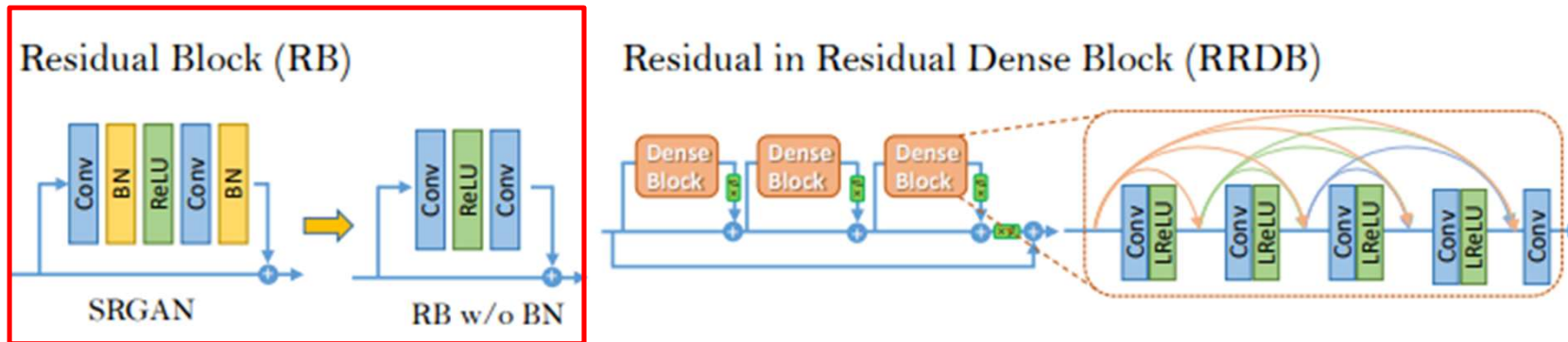


○ 본 논문에서는 대부분의 계산이 LR feature 공간에서 수행되는 SRResNet의 기본 구조를 사용한다.

- 위 구조에서 Basic Block 에 Residual block을 사용하면 SRGAN이다.

- 성능 향상을 위해 Basic block에 RRDB (Residual-in-Residual Dense Block)를 설계하고 Batch normalization layer를 제거하였다.

Proposed method



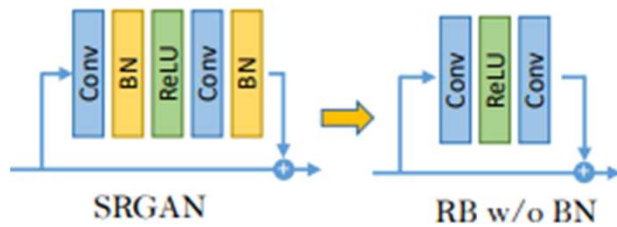
○ Batch normalization layer를 제거하면 SR 및 Deblurring을 포함한 PSNR 지향 작업에서 성능이 향상되고 계산 복잡도가 감소한다는 것이 실험적으로 입증되었다.

- BN 계층은 train 중 배치의 평균과 분산을 사용하여 feature 를 정규화 하기 때문에 train set와 test set의 통계가 많이 다를 때 아티팩트가 발생할 확률이 크다.

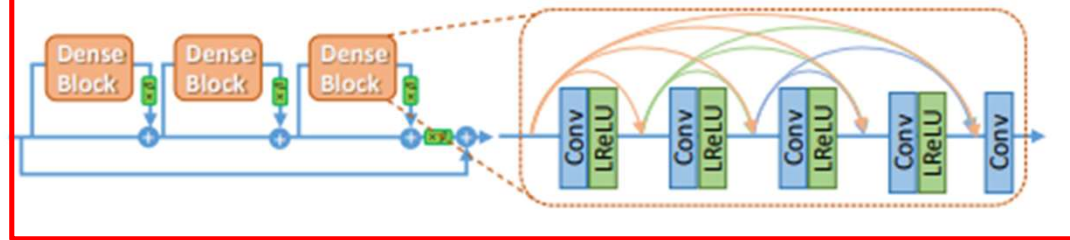
- 게다가 네트워크가 더 깊어질 수록 아티팩트가 발생할 확률이 더 커진다. 따라서 안정적인 train 과 일관된 성능을 위해 BN layer를 제거한다.

Proposed method

Residual Block (RB)



Residual in Residual Dense Block (RRDB)

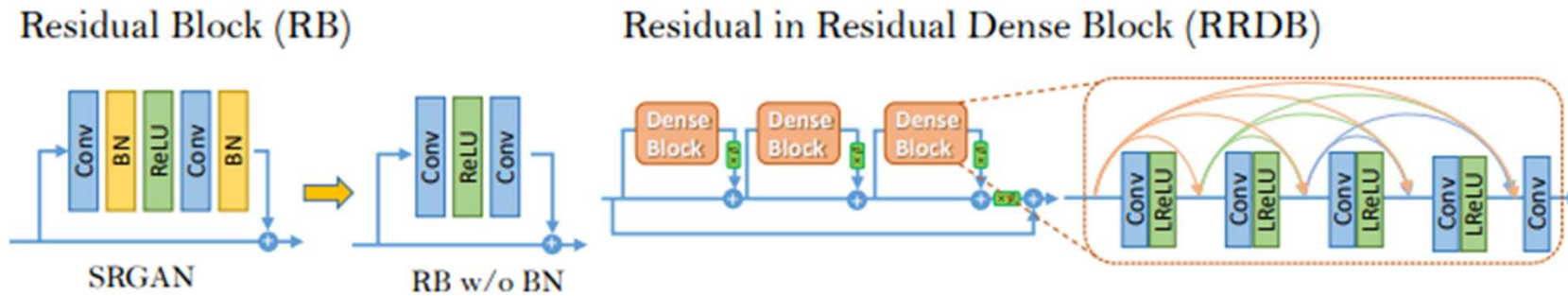


○ 본 논문에서는 SRGAN의 높은 수준의 아키텍처 설계를 유지하고 새로운 기본 블록인 RRDB를 사용한다.

- 더 많은 계층과의 연결이 항상 성능을 향상시킬 수 있다는 생각에서 제안된 RRDB는 SRGAN의 원래 residual block보다 더 깊고 복잡한 구조를 사용한다.

- train이 진행되는 경로에서 Dense block을 사용한다는 점이 특징인데 이로 인해 네트워크의 용량이 높아져 성능을 향상시킨다.

Proposed method



○ 아키텍처 개선 외에도 매우 깊은 네트워크를 훈련시키기 위한 몇가지 기술을 활용한다.

1. Residual scaling 즉 불안정성을 방지하기 위해 주 경로에 추가하기 전에 residual에 0과 1사이의 상수를 곱하여 residual을 축소한다.
2. Smaller initialization, 경험적으로 initial parameter variance가 작아질 때 residual architecture를 훈련시키기 더 쉽다는 것을 발견했다.

Proposed method

o Relativistic Discriminator

$$D(x_r) = \sigma(C(\text{Real})) \rightarrow 1 \text{ Real?}$$

$$D(x_f) = \sigma(C(\text{Fake})) \rightarrow 0 \text{ Fake?}$$

a) Standard GAN



$$D_{Ra}(x_r, x_f) = \sigma(C(\text{Real}) - \mathbb{E}[C(\text{Fake})]) \rightarrow 1 \text{ More realistic than fake data?}$$

$$D_{Ra}(x_f, x_r) = \sigma(C(\text{Fake}) - \mathbb{E}[C(\text{Real})]) \rightarrow 0 \text{ Less realistic than real data?}$$

b) Relativistic GAN

- Standard GAN의 discriminator는 하나의 입력 이미지가 실제이고 자연스러울 확률을 추정
- Relativistic discriminator는 실제 이미지 x_r 이 가짜 이미지 x_f 보다 상대적으로 더 사실적일 확률을 예측하려고 한다.
- x_f 는 input 이미지를 generator을 입력으로 넣었을 때 얻는 결과 이미지이다.

Proposed method

- Adversarial loss

$$L_G^{Ra} = -\mathbb{E}_{x_r}[\log(1 - D_{Ra}(x_r, x_f))] - \mathbb{E}_{x_f}[\log(D_{Ra}(x_f, x_r))],$$

- 새로운 Generator에 대한 adversarial loss는 실제 이미지와 생성된 이미지를 모두 포함하여 loss를 계산한다.
- 기존의 adversarial loss는 생성된 이미지만을 사용하여 loss를 계산한다.
- 따라서 본 논문의 generator는 adversarial train에서 생성된 데이터와 실제 데이터 모두에서 정보를 얻어 loss를 계산하기 때문에 성능이 높게 나온다는 것을 실험적으로 확인하였다.

Proposed method

- Perceptual loss

$$L_G = L_{\text{percep}} + \lambda L_G^{Ra} + \eta L_1$$

- Activation 한 후의 feature 가 아니라 전의 feature에 제약을 가함으로써 보다 효과적인 perceptual loss를 제안하여 다음과 같은 두가지 단점을 개선한다.
- activation 한 후의 feature는 네트워크가 깊어질 수록 매우 희박해진다.
- activated feature를 사용하면 GT 이미지와 비교하여 reconstruct된 밝기가 일정하지 않게 된다.

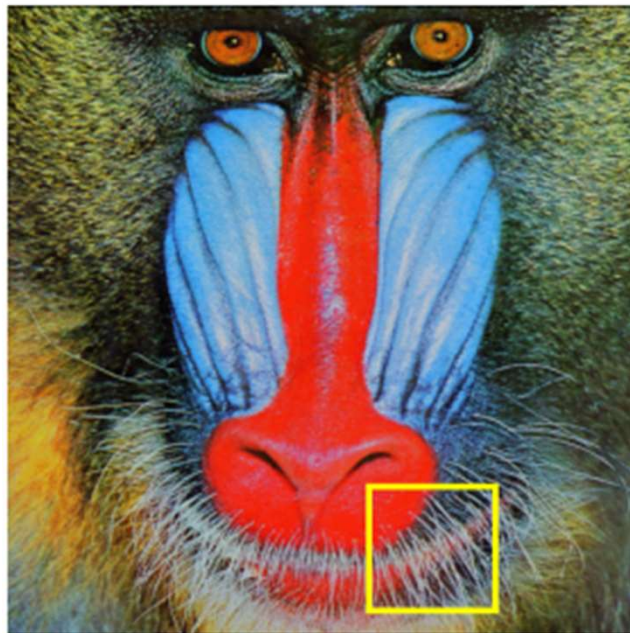
3. Experiments

2022.05.30 202132032 김희범

Experiments

- DIV2K dataset을 사용하여 진행하였다.
- 우리의 모델의 결과를 기존 PSNR-oriented 모델인 SRCNN,EDSR, RCAN 그리고 perceptual-driven 모델인 SRGAN, EnhanceNet과 비교하였다.
- perceptual 품질에 대한 효과적이고 표준적인 지표가 없기 때문에 정성적인 지표로 PSNR과 PHRM-SR 챌린지에 사용된 지각 지표(낮을 수록 좋은 결과)도 함께 제공하였다.

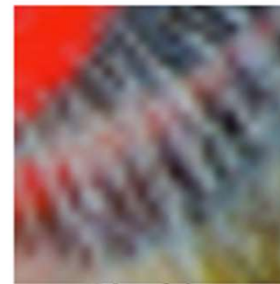
Experiments



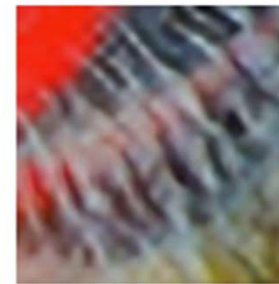
baboon from Set 14
(PSNR / Perceptual Index)



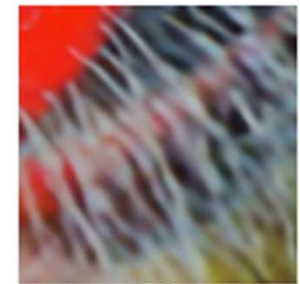
HR
(∞ / 3.59)



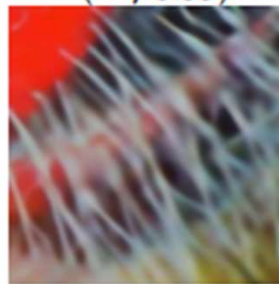
Bicubic
(22.44 / 6.70)



SRCNN
(22.73 / 5.73)



EDSR
(23.04 / 4.89)



RCAN
(23.12 / 4.20)



EnhanceNet
(20.87 / 2.68)



SRGAN
(21.15 / 2.62)



ESRGAN(ours)
(20.35 / 1.98)

Experiments



face from Set14
(PSNR / Perceptual Index)



HR
(∞ / 5.82)



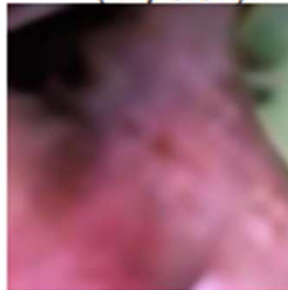
Bicubic
(31.49 / 8.37)



SRCNN
(32.33 / 6.84)



EDSR
(32.82 / 6.31)



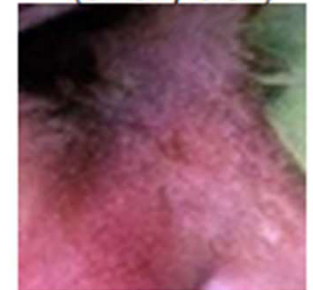
RCAN
(32.93 / 6.89)



EnhanceNet
(30.33 / 3.60)

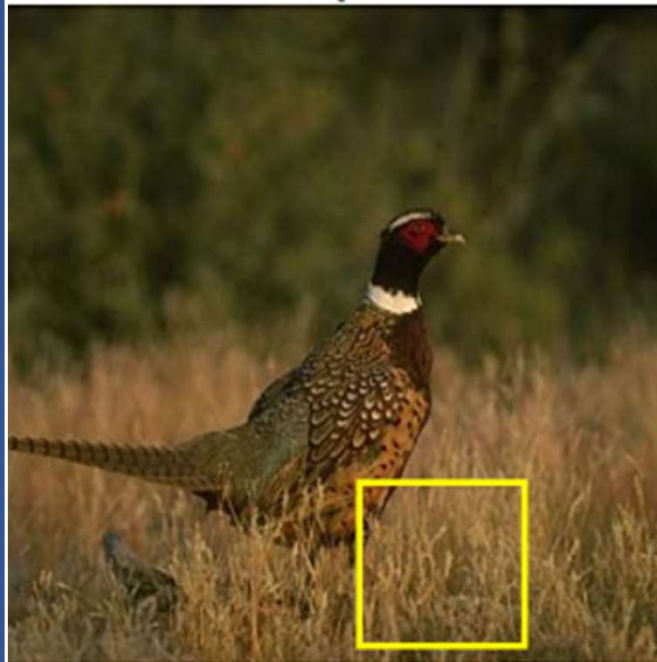


SRGAN
(30.28 / 4.47)



ESRGAN(ours)
(30.50 / 3.64)

Experiments



43074 from BSD100
(PSNR / Perceptual Index)



HR
(∞ / 2.31)



Bicubic
(29.29 / 7.35)



SRCNN
(29.62 / 6.46)



EDSR
(29.76 / 6.25)



RCAN
(29.79 / 6.22)



EnhanceNet
(27.69 / 3.00)



SRGAN
(27.29 / 2.74)



ESRGAN(ours)
(27.69 / 2.76)

4. Conclusion

2022.05.30 202132032 김희범

Conclusion

○ 기존의 PSNR oriented 모델들보다 더 나은 perceptual 품질을 달성하는 ESRGAN 모델을 제시했다.

- BN layer를 제거하고 RDDB block을 basic block으로 채택해 성능을 높였다.
- 심층 모델의 훈련을 용이하게 하기 위해 residual scaling과 smaller initialization을 사용했다.
- Relativistic Discriminator를 사용하여 한 이미지가 다른 이미지보다 더 사실적인지 여부를 판단하는 방법을 학습하여 generator가 더 자세한 texture를 복구하도록 했다.
- 결과 이미지가 다른 모델들의 결과 이미지보다 시각적으로 품질이 좋다는 것을 확인했다.

Conclusion

감사합니다