GAN-based Single Image Super Resolution

Student: Souhaiel BenSalem

20/02/2023

Background

Low-resolution image

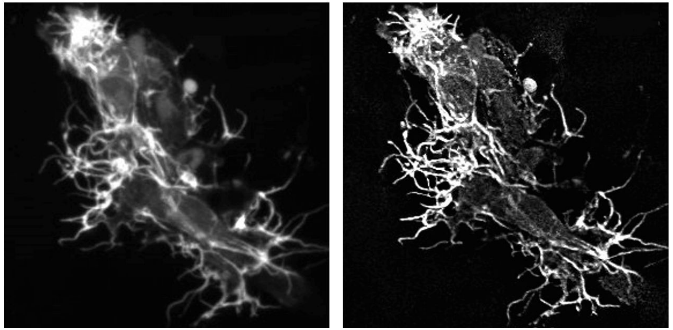
High-resolution image

SISR: the task of generating a high-resolution image from a single low-resolution image input, with the aim of retrieving the missing high-frequency details and enhancing its overall quality.

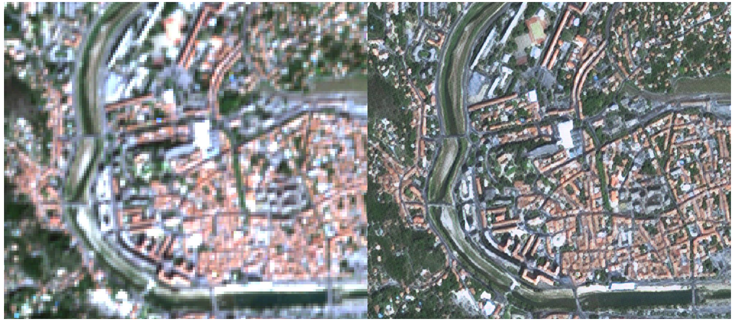
- \+ details and sharpness
- High frequency restoration
- Improved overall quality and more Creative control

Background

Medical Imaging



Remote sensing



Art restoration

Surveillance

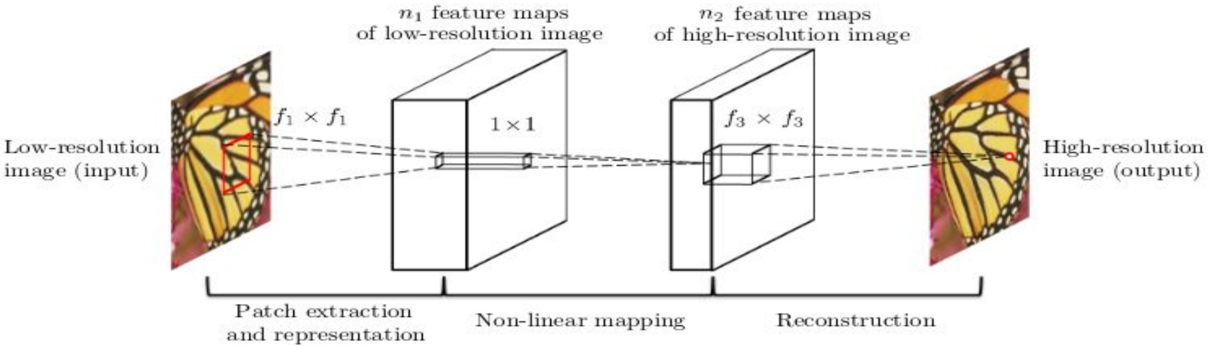
SR methods

Reconstruction based SR

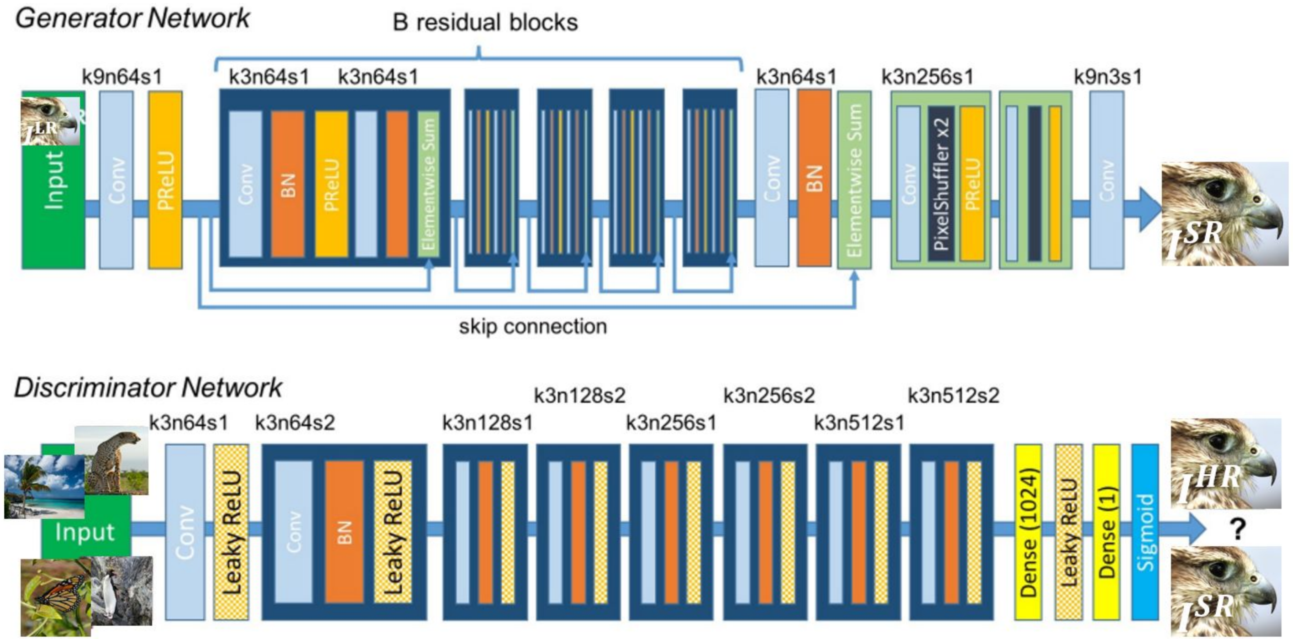
Interpolation based SR

(a) LR image Y (b) SR image X_l

Patch based SR



SRGAN



SRGAN architecture (Ledig et al 2017)

SRGAN

Authors' loss function

$$l^{SR} = \underbrace{6.10^{-3} l^{SR}_{VGG/5,4}}_{content\ loss} + \underbrace{10^{-3} l^{SR}_{Gen}}_{adversarial\ loss}$$

Our loss function

$$l^{SR}_{modified} = \underbrace{l^{SR}_{MSE}}_{pixel\ loss} + \underbrace{6.10^{-3} l^{SR}_{VGG/5,4}}_{content\ loss} + \underbrace{10^{-3} l^{SR}_{Gen}}_{adversarial\ loss} + \underbrace{2.10^{-8} l^{SR}_{TV}}_{total\ variation\ loss}$$

Methodology

Our loss function

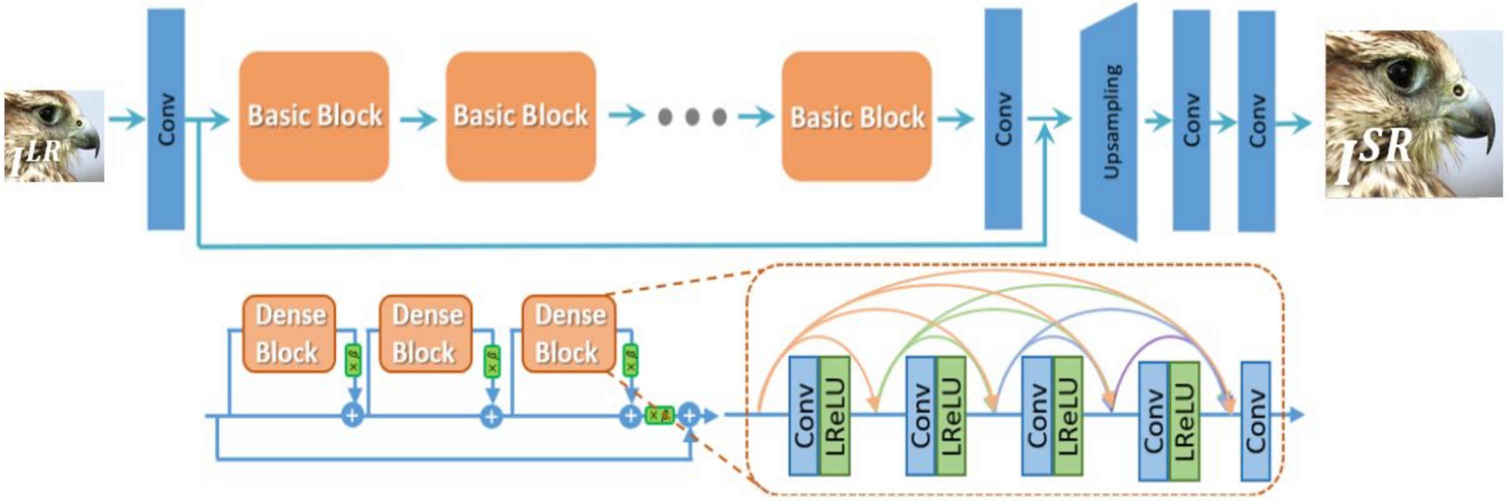
$$l^{SR}_{modified} = \underbrace{l^{SR}_{MSE}}_{pixel\ loss} + \underbrace{6.10^{-3} l^{SR}_{VGG/5,4}}_{content\ loss} + \underbrace{10^{-3} l^{SR}_{Gen}}_{adversarial\ loss} + \underbrace{2.10^{-8} l^{SR}_{TV}}_{total\ variation\ loss}$$

Intituition:

- We introduce MSE loss to penalize the differences in pixel space which ultimately leads to more accurate color fidelity between I^{SR} and I^{HR}.
- Inspired by style transfer GANs, we introduce TV loss to reduce noise in I^{SR} and use a very low weight for this loss component to avoid crashing textures in the process.

Pre-train the Generator using an MAE loss function:
Instead of pre-training the Generator network using MSE loss, we used experimented with a **weighted MSE** and $L1$ **loss**, which both yielded better results than MSE from a human viewer's perspective, and settled for the latter.

ESRGAN



ESRGAN Generator network (Wang et al 2018)

ESRGAN

ESRGAN:

- improves on the ideas introduced by SRGAN to achieve a better perceptual quality.
- The improvements are mainly technical and consist of modifying the architecture of the generator network.
- The authors introduced the a residual in residual block and removed batch normalization layers.

The loss function was also modified by introducing L_1 penalty and taking VGG features before activation for the perceptual loss component.

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

Evaluation and Results

The models were trained on the DIV2K dataset that contains:

- 1600 training images of different sizes divided into 800 high resolution images and their corresponding x4 down-scaled and bicubic-interpolated low resolution images.
- 200 test images divided into 100 high resolution images and their corresponding low resolution counterparts.

The models were also evaluated on the Set5 and Set14 benchmark datasets.

Evaluation and Results

Pixel burning artifacts while training SRGAN, also known as BatchNorm artifacts

Evaluation and Results

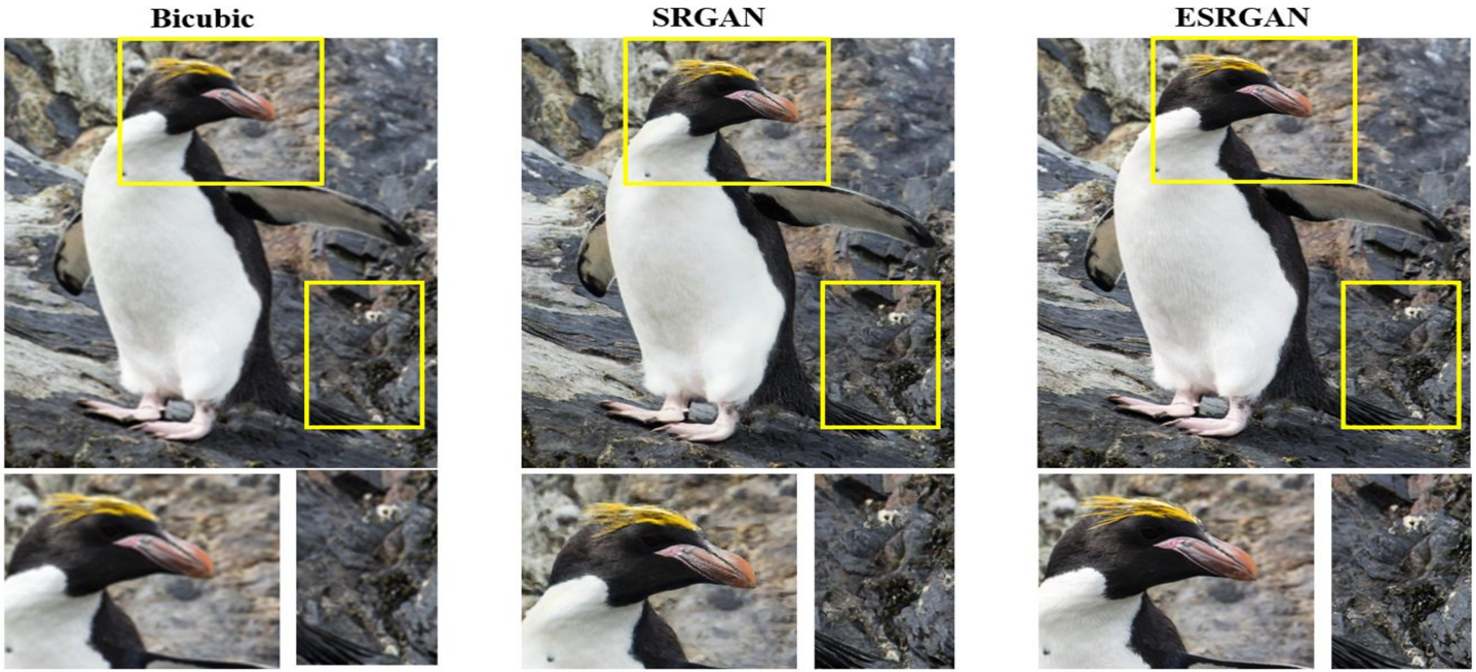
DIV2K	SRGAN(ours)	SRGAN (Ledig et al)	ESRGAN
PSNR	25.373	-	28.174
SSIM	0.706	-	0.775

Set5			
PSNR	24.945	29.40	30.474
SSIM	0.718	0.8472	0.851

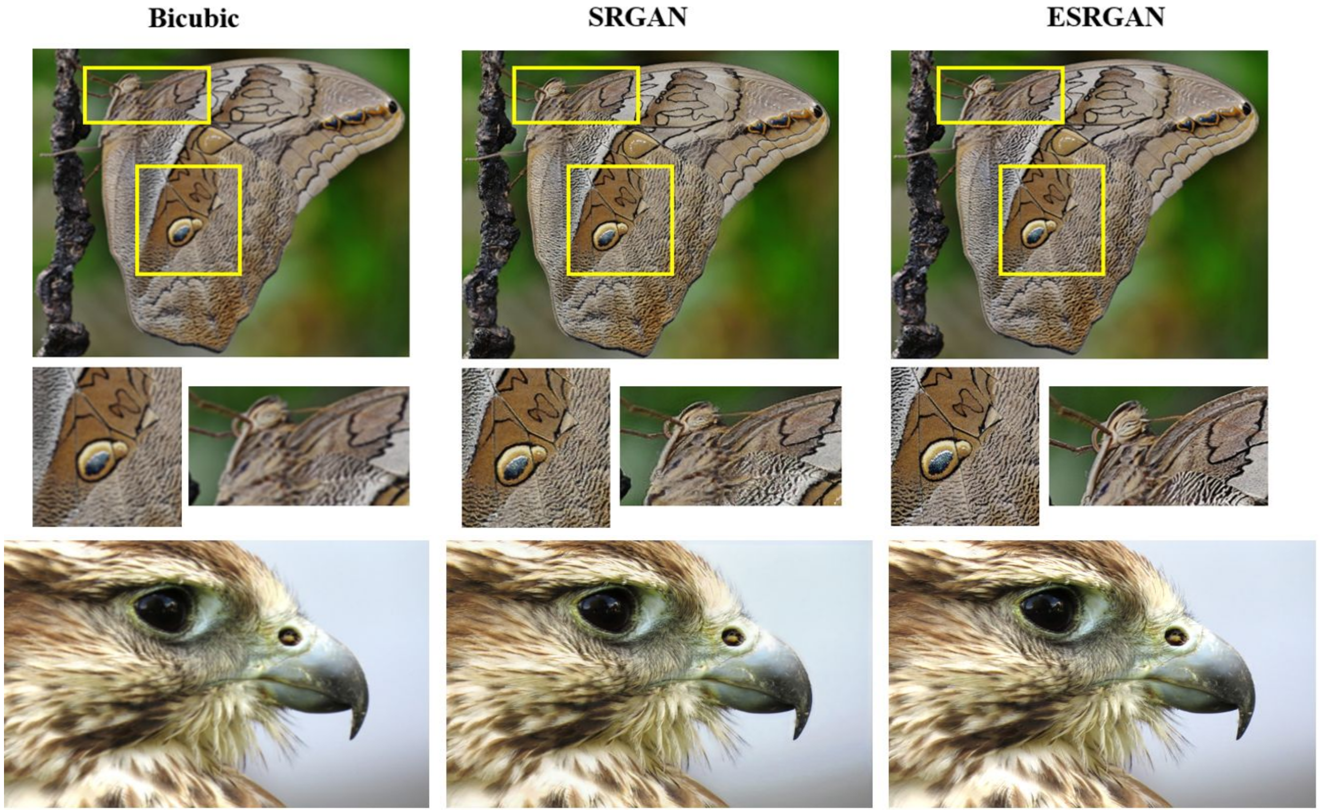
Set14			
PSNR	23.770	26.02	26.614
SSIM	0.636	0.7397	0.713

Results after 850 epochs of training

Evaluation and Results



Evaluation and Results



Thank You !