Applications of GANs

- Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network

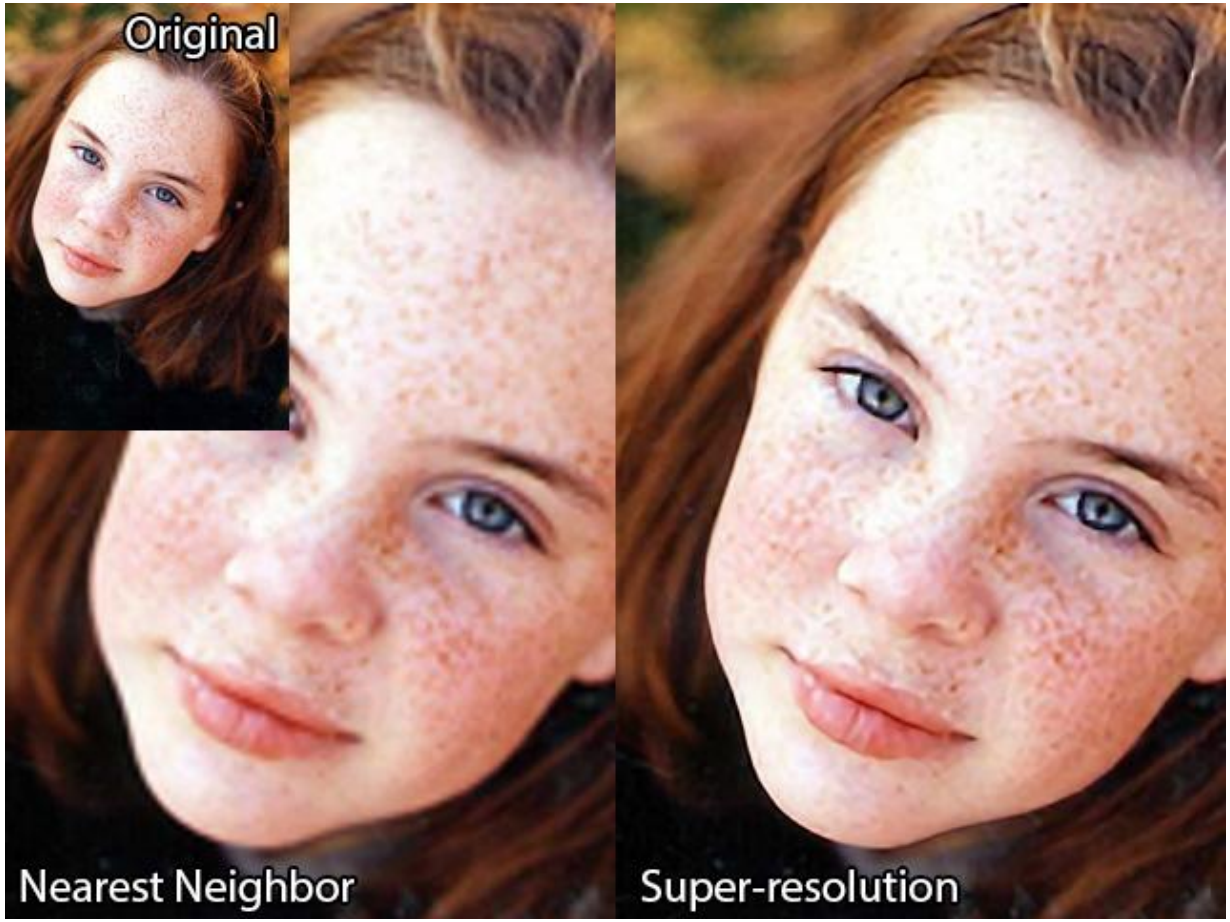
Using GANs for Single Image Super-Resolution

Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, Wenzhe Shi

Problem

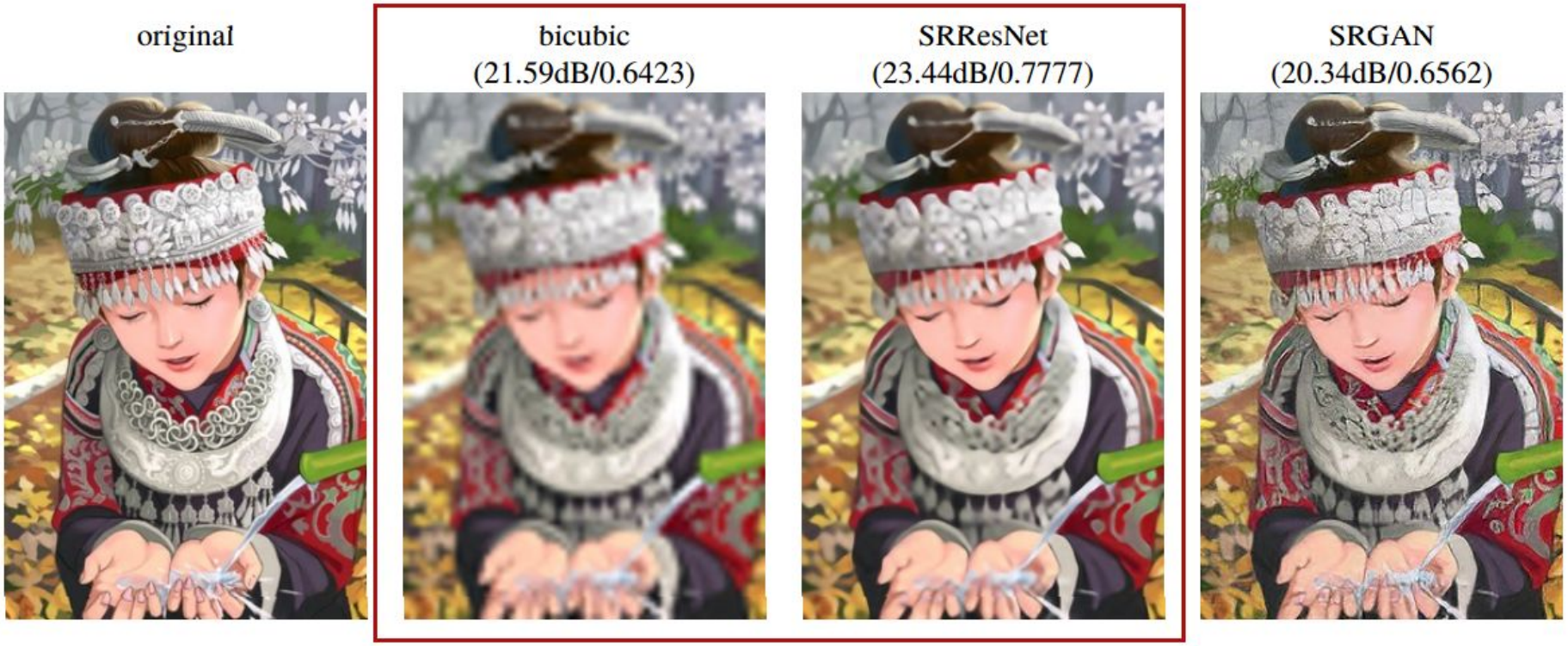
How do we get a high resolution (HR) image from just one (LR) lower resolution image?

Answer: We use super-resolution (SR) techniques.

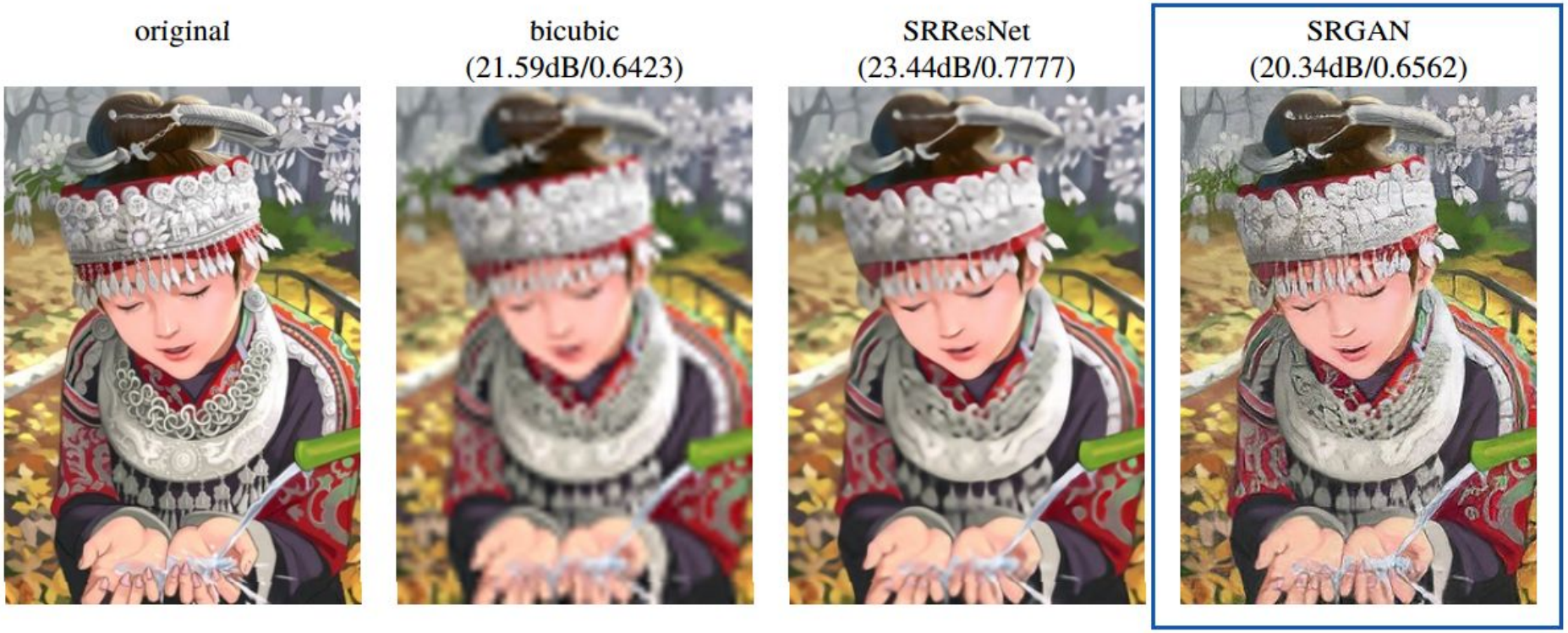


http://www.extremetech.com/wp-content/uploads/2012/07/super-resolution-freckles.jpg

Previous Attempts



SRGAN


original
bicubic
(21.59dB/0.6423)
SRResNet
(23.44dB/0.7777)
SRGAN
(20.34dB/0.6562)

SRGAN - Generator

- G: generator that takes a low-res image I^{LR} and outputs its high-res counterpart I^{SR}
- θ_G: parameters of G, $\{W_{1:L}, b_{1:L}\}$
- l^{SR}: loss function measures the difference between the 2 high-res images

$$\hat{\theta}_G = \arg\min_{\theta_G} \frac{1}{N} \sum_{n=1}^{N} l^{SR}(G_{\theta_G}(I_n^{LR}), I_n^{HR})$$

SRGAN - Discriminator

- D: discriminator that classifies whether a high-res image is I^{HR} or I^{SR}
- θ_D: parameters of D

$$\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{I^{HR} \sim p_{\text{train}}(I^{HR})} [\log D_{\theta_D}(I^{HR})] + \mathbb{E}_{I^{LR} \sim p_G(I^{LR})} [\log(1 - D_{\theta_D}(G_{\theta_G}(I^{LR})))]$$

SRGAN - Perceptual Loss Function

Loss is calculated as weighted combination of:

- ➔ Content loss
- ➔ Adversarial loss
- ➔ Regularization loss

SRGAN - Content Loss

Instead of MSE, use loss function based on ReLU layers of pre-trained VGG network. Ensures similarity of content.

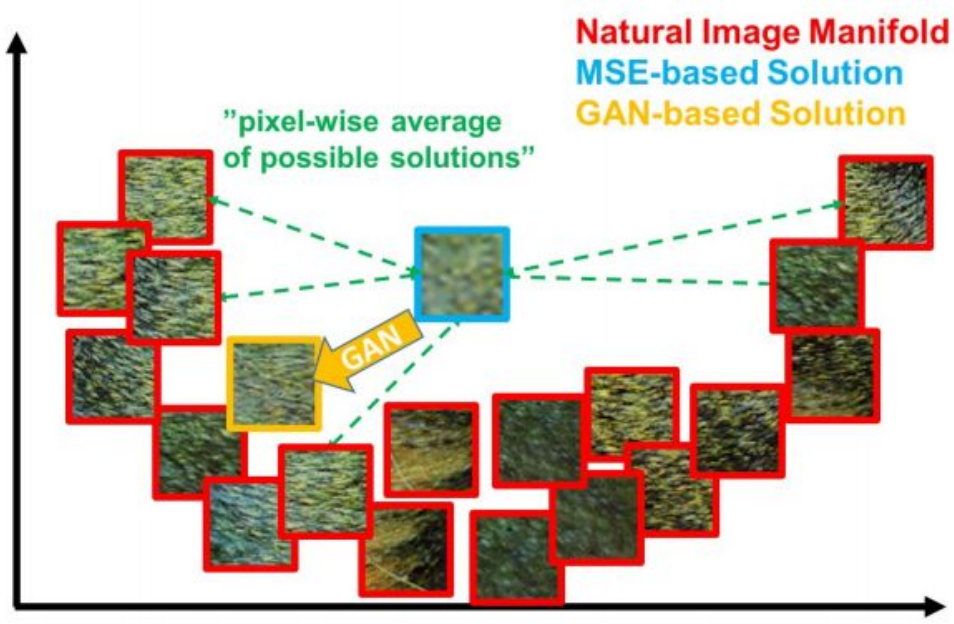
- $\boldsymbol{\Phi}_{i,j}$: feature map of j^{th} convolution before i^{th} maxpooling
- $W_{i,j}$ and $H_{i,j}$: dimensions of feature maps in the VGG

$$l^{SR}_{VGG/i.j} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

SRGAN - Adversarial Loss

Encourages network to favour images that reside in manifold of natural images.

$$l_{Gen}^{SR} = \sum_{n=1}^{N} -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$



SRGAN - Regularization Loss

Encourages spatially coherent solutions based on total variations.

$$l_{TV}^{SR} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} ||\nabla G_{\theta_G}(I^{LR})_{x,y}||$$

SRGAN - Examples

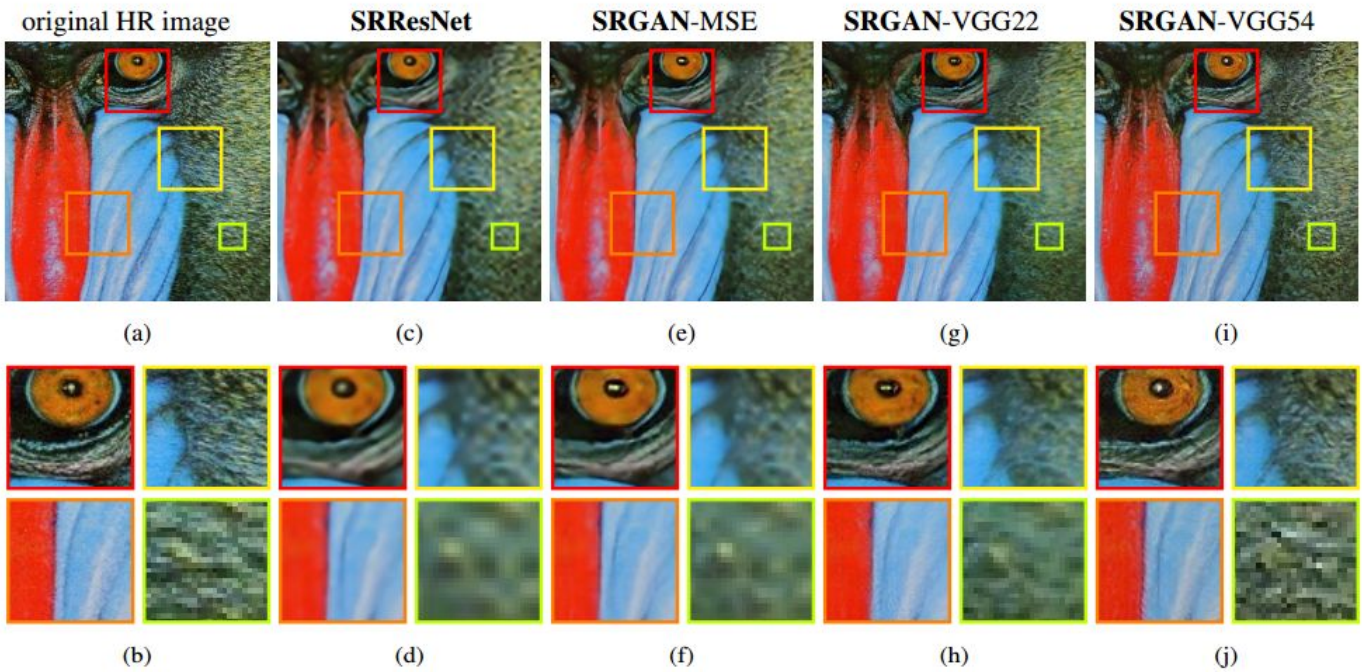


Figure 5: Reference HR image (left: a,b) with corresponding SRResNet (middle left: c,d), SRGAN-MSE (middle: e,f), SRGAN-VGG2.2 (middle right: g,h) and SRGAN-VGG54 (right: i,j) reconstruction results.

SRGAN - Examples