Image Super–Resolution with GANs

Nathanael Bosch[1], Thomas Grassinger[1], Jonas Kipfstuhl[1], Pierre Springer[1]

[1]Technical University of Munich

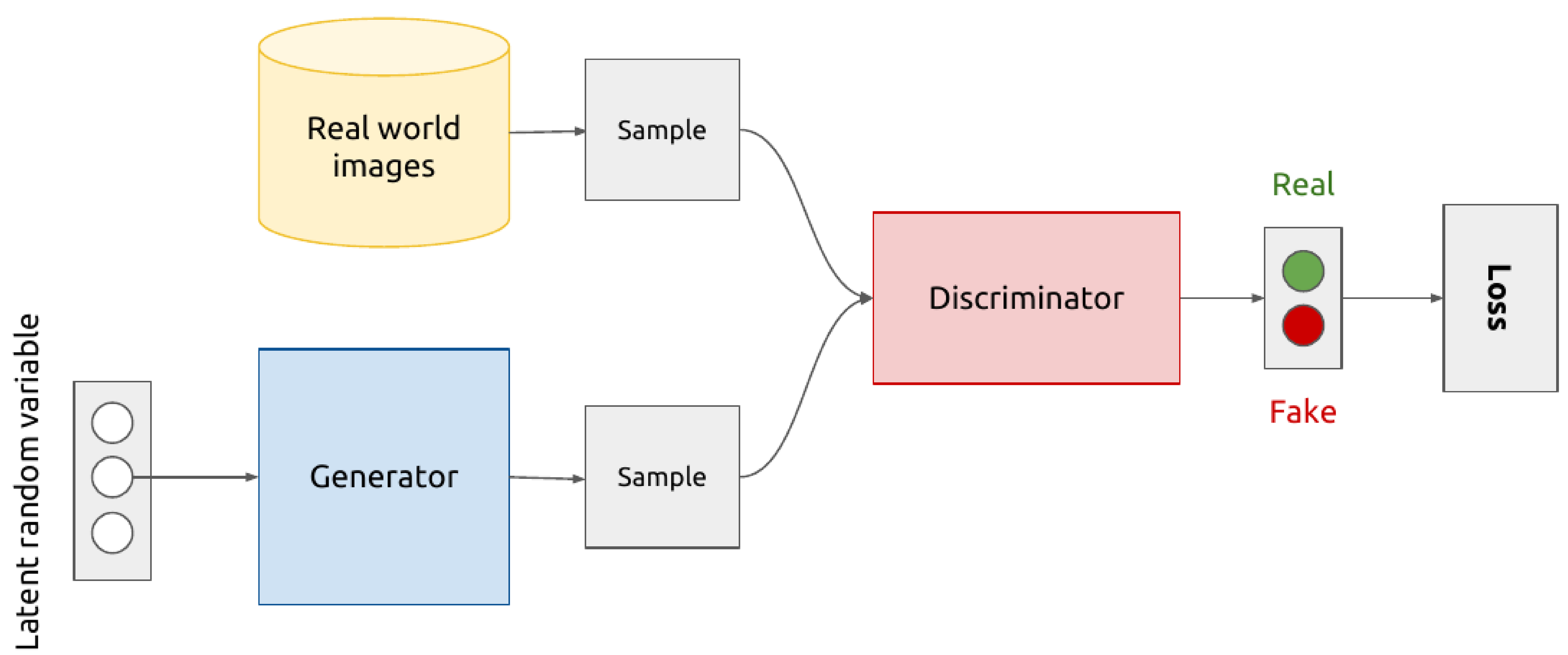
Introduction

We call super-resolution (SR) the task of estimating a high-resolution (HR) image from its low-resolution (LR) counterpart.
Recent work with optmizition-based methods largely focuses on minimizing the mean squared reconstruction error. This results in high peak signal-to-noise ratios, but they often have problems with modelling high-frequency details. The results are often too smooth. In our work we tried to tackle this problem using generative adversarial networks (GANs).

GANs

GANs consist of two different networks, a Generator Network and a Discriminator Network. The concept behind this is that the generative network estimates a super-resolved image from its LR version with the goal to become highly similar to real images that the discriminator network fails to distiguish. Netword Architektur einfügen
Therefore we optimize the discriminator network D_{Θ_D} in an alternating manner along with the generative network G_{Θ_G} to solve the adversarial min-max problem:

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



The perceputal loss l^{SR} we defined as weighted sum of a content loss and an discriminative loss component:

$$l^{SR} = \alpha l^{SR}_{MSE} + \beta l^{SR}_{VGG16_19/i.j} + \gamma l^{SR}_{D}$$

Where l^{SR}_{MSE} is a mean squared error term, l^{SR}_{VGG} is an euclidean distance of the VGG feature representation for the images, and l^{SR}_{D} is the discriminative loss.

Datasets

PASCAL VOC[1] over 10 000 images for classification in 20 categories

NITRE[2] 800 high resolution images

Results

(a) VGG1619 perceptual, adversarial, image loss

(b) VGG1619 perceptual, adversarial loss

(c) VGG1619 perceptual, image loss

(d) VGG1619 perceptual loss

(e) VGG1619 adversarial, image loss

(a) VGG1619 adversarial loss

(b) VGG1619 image loss

(c) VGG19 perception, adversarial, image loss

(d) VGG16 perception, adversarial, image loss

Key Points

- Networks learn even without discriminator
- VGG1619 performs considerably better than VGG16
- Training works without discriminator
- The discriminator is hard to train
- Image metrics don't change much, but the images still improve

[1]http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html
[2]https://data.vision.ee.ethz.ch/cvl/DIV2K/