DCGAN Image Patcher

Image bits of images to a fully-formed image

November 17, 2024

The Standard Game

$$\min_{\theta_G} J^{(G)}(\theta_D, \theta_G) = \min_{\theta_G} \mathbb{E}_z[\log(1 - D(G(z)))]$$

$$\max_{\theta_D} J^{(D)}(\theta_D, \theta_G) = \max_{\theta_D} \left(\mathbb{E}_x \left[\log(D(x))\right] + \mathbb{E}_z \left[\log(1 - D(G(z)))\right]\right)$$

Our Modified Game

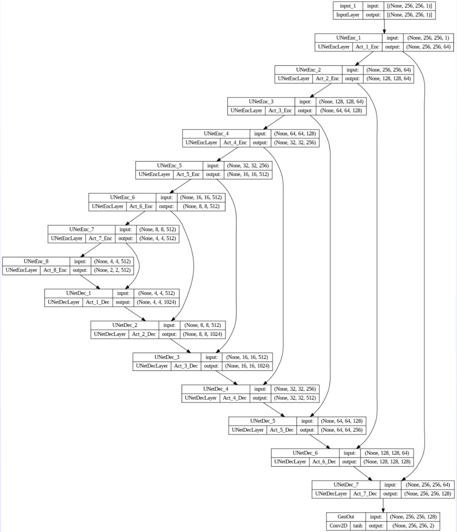
- Generator Learning Stagnation:
  If the discriminator is too effective, the generator's gradients can nearly vanish during training, leading to repeated similar outputs and slow learning progress.
- Cost Function Divergence:
  The original GAN cost function can potentially decrease indefinitely, risking divergence to negative infinity and disrupting stable network training. (Goodfellow)

$$\max_{\theta_G} J^{(G)^*}(\theta_D, \theta_G) = \max_{\theta_G} \mathbb{E}_z \left[\log(D(G(z)))\right]$$

$$\min_{\theta_G} J^{(G)*}(\theta_D, \theta_G) = \min_{\theta_G} -\mathbb{E}_z[\log(D(G(z)))] + \lambda \|G(z) - y\|_1$$

Aim of l^1 regularization: ensure the preservation of the original images' structure. and prevent the generator from manipulating colors just to mislead the discriminator.

U-Net Architecture for Generator



Tweaks

- One-sided label smoothing: encourage the discriminator to estimate soft probabilities and reduce vulnerability of GAN to adversarial examples (Saliman et. al.)
- Reduced momentum: reduce instability in training (Alec Radford et. al.)
- LAB encoding instead of RGB: small perturbations in LAB do not change the image too much unlike in RGB

Results

Bad Image	Generated Image	Actual Image
		
		
		

References

- Ian Goodfellow. Nips 2016 tutorial: Generative adversarial networks. 2016.
- Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. 2015
- Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. In Advances in Neural Information Processing Systems, pages 2234–2242, 2016.