Figure: Randomly generated images using QAGANs for CelebA, STL10 datasets. Right : QAGAN- SSIM , Left: QAGAN-NIQE.

Quality Aware Generative Adversarial Networks

Parimala Kancharla , Sumohana S. Channappayya
Department of Electrical Engineering, Indian Institute of Technology Hyderabad. India
{ee15m17p100001,sumohana@iith.ac.in}

NeurIPS 2019

Objective and approach

- Goal: To design quality aware regularizers for training Generative Adversarial Networks.
- Approach: **Using the full reference IQA algorithm (SSIM) and No reference IQA algorithm (NIQE).**

Background

- Generative Adversarial Networks (GANs) [1] are generative models designed to learn the probability distribution of data that is aided by adversarial learning.
- A GAN is composed of two models: The generator model $G(z;\theta_G)$ the discriminator model $D(x;\theta_D)$
- **Objective Function:**

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log(D(x))] + E_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

Wasserstein Generative Adversarial Network

$$W(p_r, p_\theta) = \max_{w\epsilon W} E_{x\sim p_r}[D_w(x)] - E_{z\sim p_z}[D_w(g_\theta(z)]$$

- Lipschitz constraint is enforced by weight clipping

$$w\epsilon[-c,c]$$

WGAN -GP:

- Gradient penalty for discriminator loss to impose the lipschitz constraint.

$$|D(x)-D(y)| \leq ||x-y||_2$$

$$L = E_{x\sim p_r}[D_w(x)] - E_{z\sim p_z}[D_w(g_\theta(z)] + \lambda E_{\hat{x}}(||\nabla D_w(\hat{x})||_2 - 1)^2$$

- Where the distribution of $\hat{x}$ is taken to be the uniform distributions onlines connecting points drawn from p_r and $p_{g_\theta(z)}$.

Structural Similarity Index (SSIM)

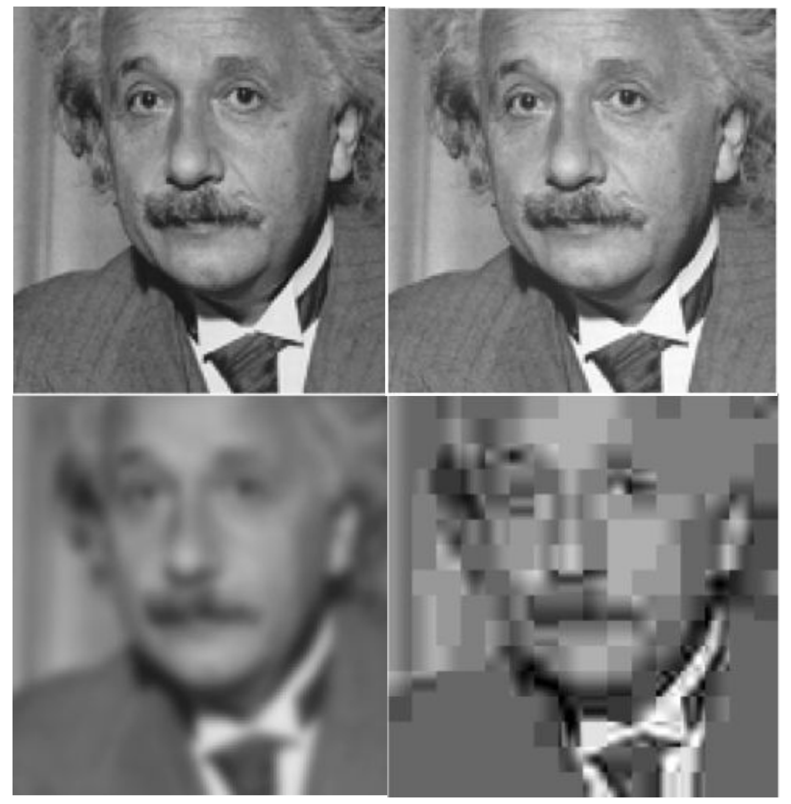

Figure: Same MSE! but different perceptual quality

QAGAN-SSIM

$$SSIM(P_{(i,j)}, T_{(i,j)}) = L(P_{(i,j)}, T_{(i,j)}).C(P_{(i,j)}, T_{(i,j)}).S(P_{(i,j)}, T_{(i,j)})$$

- Where P, T refers to the pristine and test images.
- $L(P_{(i,j)}, T_{(i,j)})$ $C(P_{(i,j)}, T_{(i,j)})$ $S(P_{(i,j)}, T_{(i,j)})$ and are the local luminance, contrast and structure scores at pixel (i, j) respectively. Further,

$$L(P_{(i,j)}, T_{(i,j)}) = \frac{2\mu_P(i,j)\mu_T(i,j)+C_1}{\mu_P^2(i,j)+\mu_T^2(,j)+C_1} \quad C(P_{(i,j)}, T_{(i,j)}) = \frac{2\sigma_P(i,j)\sigma_T(i,j)+C_2}{\sigma_P{}^2(i,j)+\sigma_T{}^2(i,j)+C_2} \quad S(P_{(i,j)}, T_{(i,j)}) = \frac{\sigma_{PT}(i,j)+C_3}{\sigma_P(i,j)\sigma_T(i,j)+C_3}$$

SSIM Metric: $d^Q(P_{(i,j)}, T_{(i,j)}) = \sqrt{2 - L(P_{(i,j)}, T_{(i,j)}) - CS(P_{(i,j)}, T_{(i,j)})}$

$$0 \leq d^Q(x,y) \leq \sqrt{2}$$

- Discriminator should be lipstichtz with respect to $d^Q(X,Y)$.

$$\text{SSIM GP} = E_{X\sim P_r, Y\sim P_G}\left[\left(\frac{|D(X)-D(Y)|}{d^Q(X,Y)}\right)-1\right]^2$$

$$L_d = \min_{D\epsilon\mathcal{D}}\left(E_{z\sim P_z}D(G(z)) - E_{X\sim P_r}D(X)\right) + \lambda_1 E_{\hat{x}\sim P_{\hat{x}}}(||\nabla_{\hat{x}}D(\hat{x})||_2-1)^2 + \lambda_2 E_{X\sim P_r, G(z)\sim P_G}\left[\left(\frac{|D(X)-D(G(z))|}{d^Q(X,G(z))}\right)-1\right]^2$$

Natural Image Quality Evaluator (NIQE)

- NIQE builds a statistical model for the class of pristine natural scenes.
- Mean subtracted contrast normalization: $\hat{I}(i,j) = \frac{I(i,j)-\mu(i,j)}{\sigma(i,j)+1}$
- The parameters of these GGD and AGGD models are in turn modeled using a Multivariate Gaussian (MVG) distribution.
- NIQE score: the "distance" of its MVG parameters from the pristine MVG parameters.

$$D(\mu_P, \mu_T, \Sigma_P, \Sigma_T) = \sqrt{(\mu_P - \mu_T)^T\left(\frac{\Sigma_P + \Sigma_T}{2}\right)^{-1}(\mu_P - \mu_T)}$$

QAGAN - NIQE

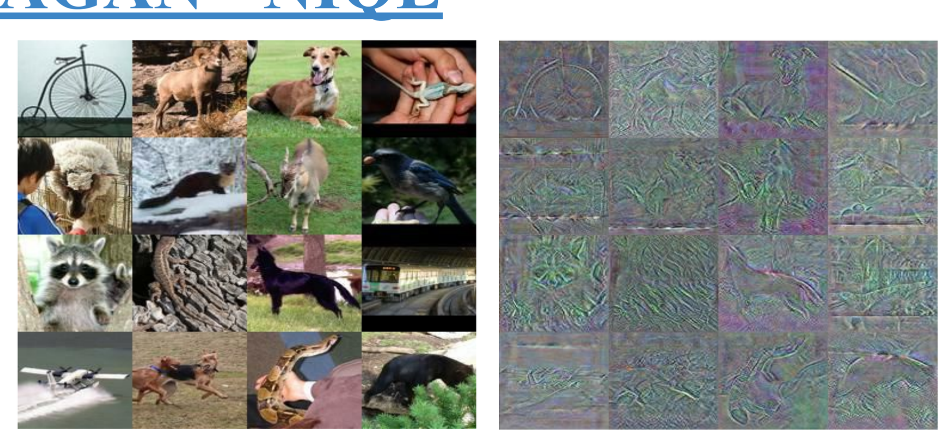

Figure: Real images and corresponding discriminator gradient maps.

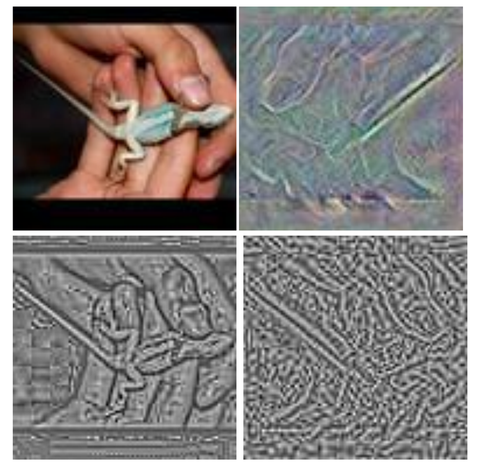

Figure: MSCN coeffiecients.

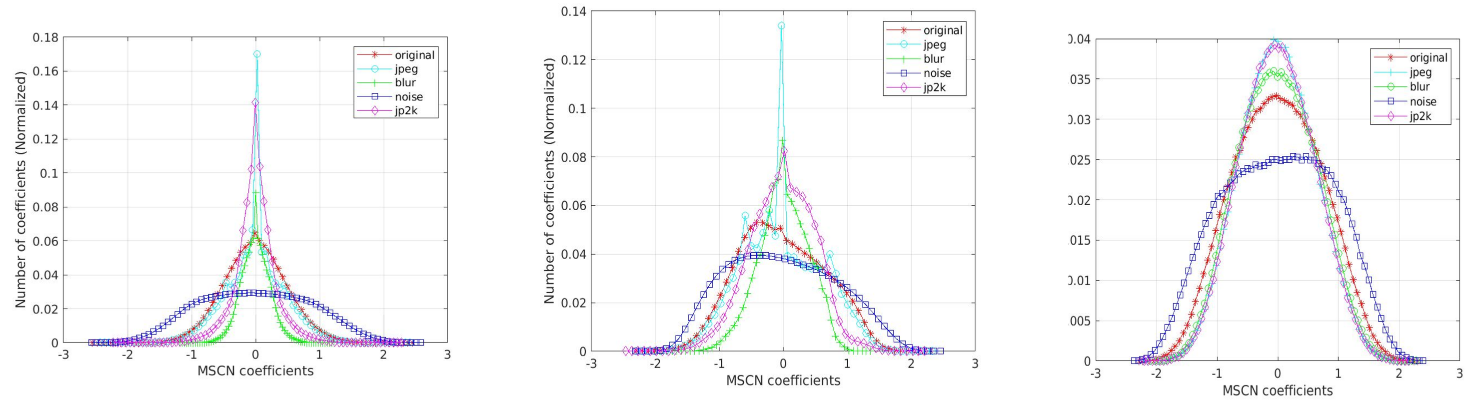

Figure: The empirical histograms of MSCN coefficients

- Motivation: To preserve the MSCN coefficients of the discriminator gradient map.
- The naturalness of a test discriminator gradient image T is computed to be its "distance" from the pristine image gradient class.

$$L_d = \min_{D\epsilon\mathcal{D}}\left(E_{z\sim P_z}D(G(z)) - E_{X\sim P_r}D(X)\right) + \lambda_1 E_{\hat{x}\sim P_{\hat{x}}}(||\nabla_{\hat{x}}D(\hat{x})||_2-1)^2 + \lambda_2 E_{\hat{x}\sim P_{\hat{x}}}(||(\nabla_{\hat{x}}D(\hat{x})|\mu_P, \Sigma_P)||_{\text{NIQE}})$$

- where λ 's are hyper parameters and those are chosen empirically.

Evaluation

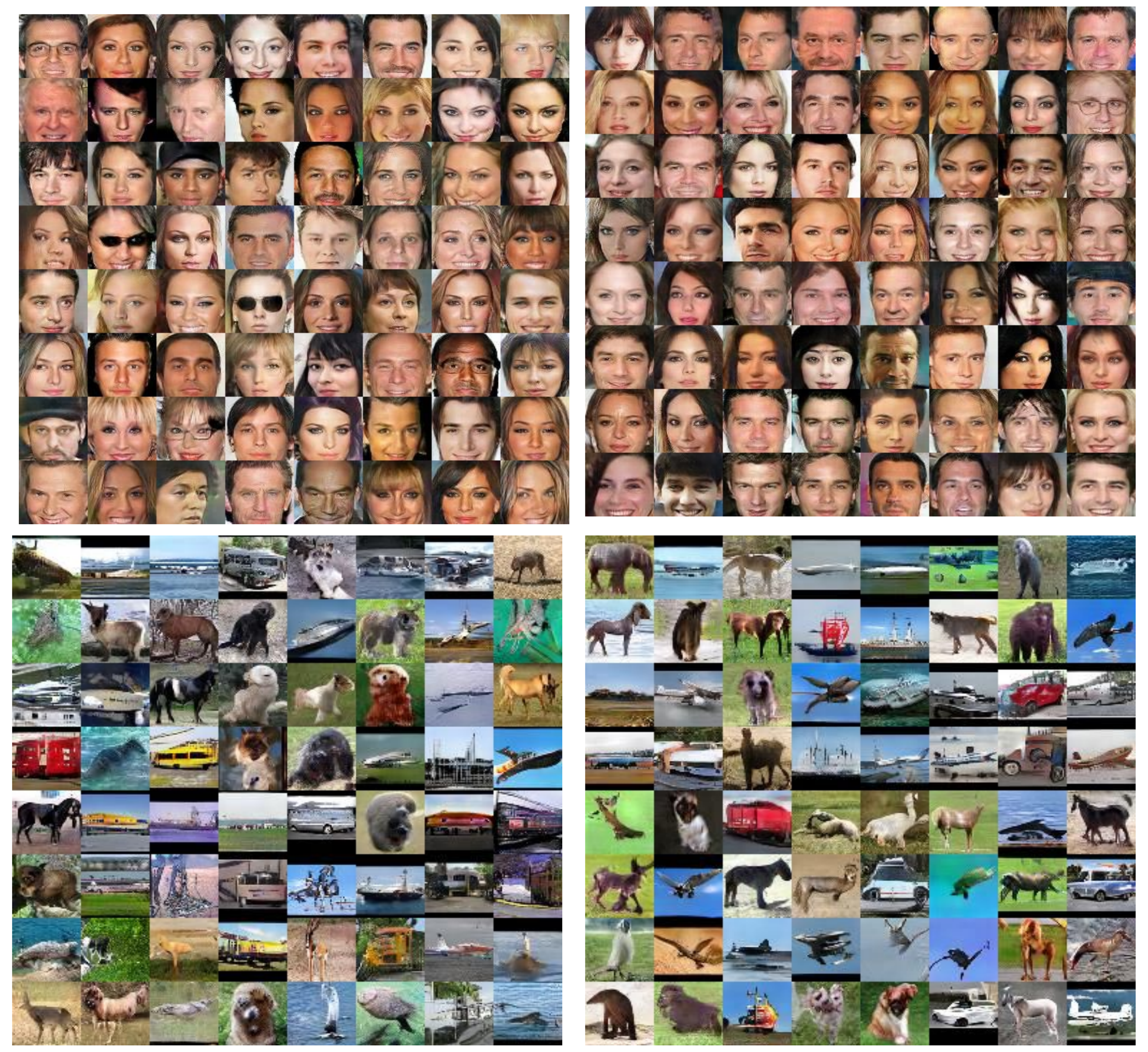

Figure: Randomly generated images using QAGANs for CelebA, STL10 datasets. Right : QAGAN- SSIM , Left: QAGAN-NIQE.

- FID and IS scores on CelebA, STL10 and CIFAR10 datasets.

Model	FID
Real Faces (CelebA)	1.09
WGAN-GP	12.89
Banach WGAN	10.5
MMD GAN-rep-b	6.79
QAGAN (SSIM)	**6.421**
QAGAN (NIQE)	**6.504**

Model	IS	FID
Real Data (48×48)	26.08 ± 0.26	7.9
WGAN-GP	9.05 ± 0.12	55.1 ± 0.0
SNGAN	9.10 ± 0.04	40.10 ± 0.50
MMD GAN-rep	9.36 ± 0.0	36.67 ± 0.0
QAGAN (SSIM)	$\mathbf{9.29 \pm 0.05}$	$\mathbf{19.77 \pm 0.0091}$
QAGAN (NIQE)	$\mathbf{9.1720 \pm 0.08}$	$\mathbf{19.45 \pm 0.0013}$

Model	IS	FID
Real data	11.24 ± 0.12	7.80
DCGAN	6.16 ± 0.07	-
WGAN-GP	7.86 ± 0.10	40.2 ± 0.0
CTGAN	8.12 ± 0.12	-
SNGAN	8.12 ± 0.12	21.5 ± 0.21
$W^{-\frac{3}{2},2}$ - Banach WGAN	8.26 ± 0.07	-
L^{10} - Banach WGAN	8.31 ± 0.07	-
MMD GAN-rep-b	8.29 ± 0.0	$16.21 \pm 0,0$
QAGAN (SSIM)	$\mathbf{8.37 \pm 0.04}$	$\mathbf{13.91 \pm 0.105}$
QAGAN (NIQE)	$\mathbf{7.87 \pm 0.027}$	$\mathbf{12.4697\pm 0.068}$

Conclusions:

- Based on insights from both FR and NR IQA metrics, we have proposed two novel regularization approaches for the WGAN-GP framework.
- The key takeaway from our work is that the unique local structural and statistical signature of pristine natural images must be preserved in the generated images.

References:

1.https://www.iith.ac.in/~lfovia/
2. Ian Goodfellow et al. Generative adversarial nets. In: Advances in neural information processing systems.
3. Martin Arjovsky, Soumith Chintala, and Lon Bottou. Wasserstein generative adver- sarial networks. In: International Conference on Machine Learning . 2017
4. Ishaan Gulrajani et al. Improved training of wasserstein gans. In: Advances in Neural Information Processing Systems . 2017, pp. 57675777.
5. Zhou Wang et al. Image quality assessment: from error visibility to structural similarity. In: IEEE transactions on image processing 13.4 (2004), pp. 600612
6. Dominique Brunet, Edward R Vrscay, and Zhou Wang. On the mathematical properties of the structural similarity index. In: IEEE Transactions on Image Processing 21.4 (2011), pp. 14881499
7. Daniel L Ruderman. The statistics of natural images. In: Network: computation in neural systems 5.4 (1994), pp. 517548
8. Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a completely blind image quality analyzer. In: IEEE Signal Processing Letters 20.3 (2013), pp. 209212