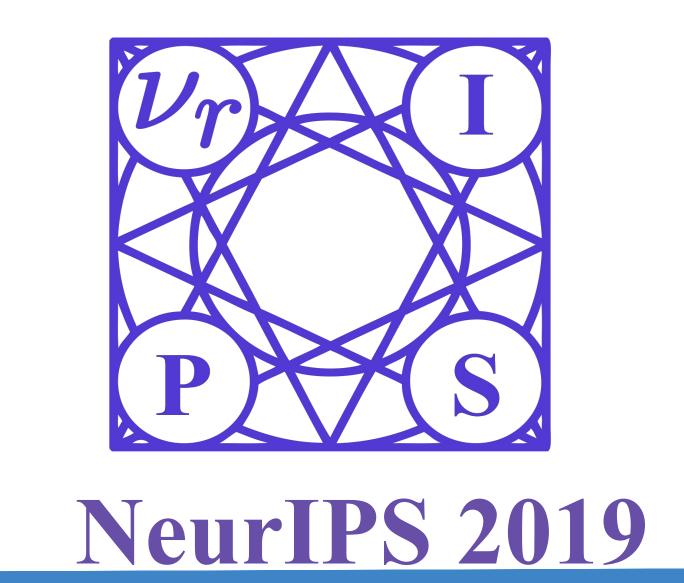




# Quality Aware Generative Adversarial Networks

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



### 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:

$$\min_{G} \max_{D} V(D,G) = E_{x \sim p_{data}(x)}[\log(D(x))] + E_{z \sim p_{z}(z)}[\log(1 - D(G(z)))] \quad \text{SSIM}(P_{(i,j)}, T_{(i,j)}) = L(P_{(i,j)}, T_{(i,j)}) \cdot C(P_{(i,j)}, T_{(i,j)}) \cdot S(P_{(i,j)}, T_{(i,j)})$$

### Wasserstein Generative Adversarial Network

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

• Lipschitz constraint is enforced by weight clipping  $w\epsilon[-c,c]$ 

#### WGAN-GP:

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

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

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

• 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)

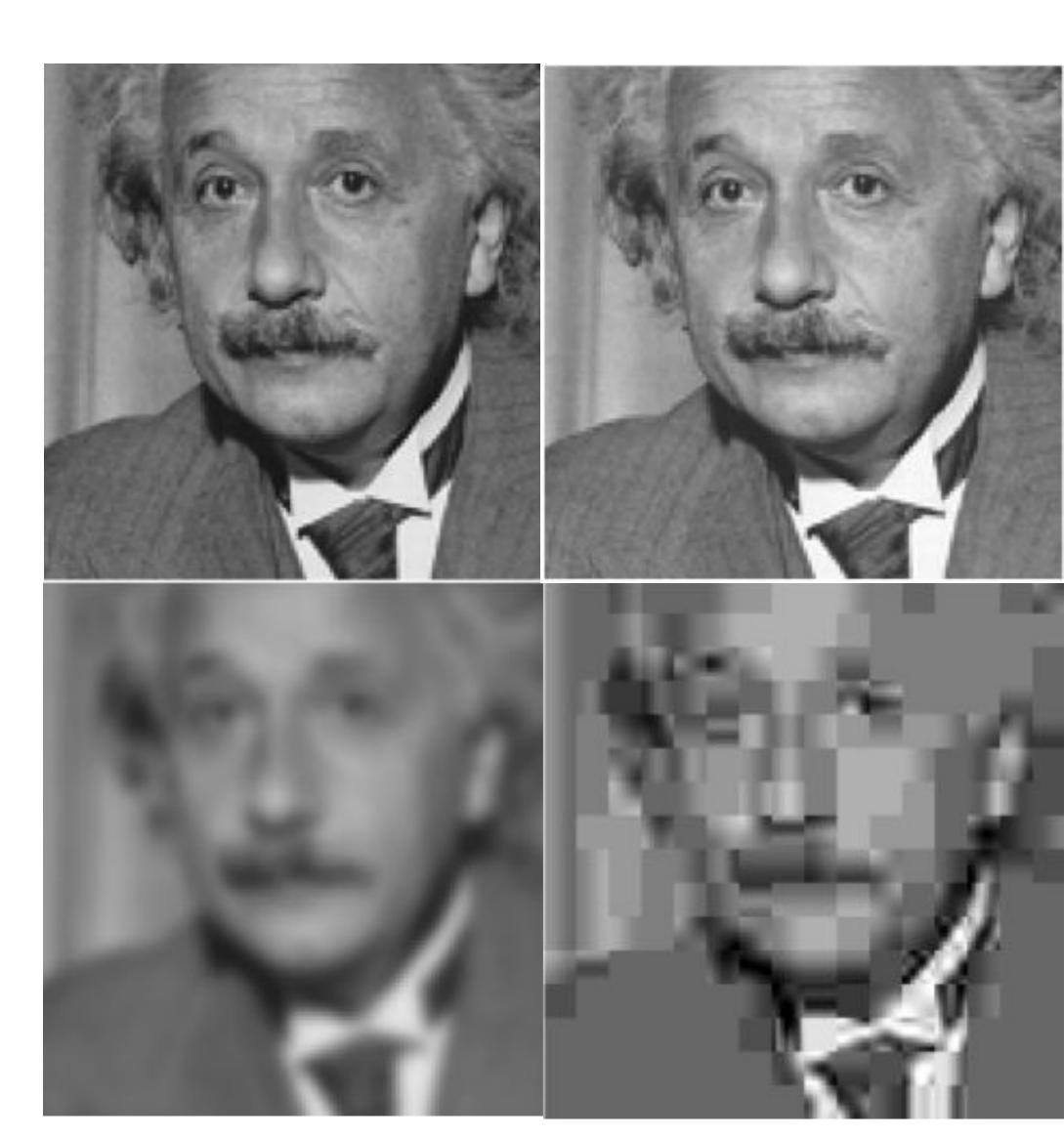


Figure: Same MSE! but different perceptual quality

### **OAGAN-SSIM**

- 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: 
$$d^Q(P_{(i,j)}, T_{(i,j)}) = \sqrt{2 - L(P_{(i,j)}, T_{(i,j)}) - CS(P_{(i,j)}, T_{(i,j)})}$$
  $0 \le d^Q(x, y) \le \sqrt{2}$ 

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

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

$$L_{d} = \min_{D \in \mathcal{D}} \left( E_{z \sim P_{z}} D(G(z)) - E_{X \sim P_{r}} D(X) \right) + \lambda_{1} E_{\hat{x} \sim P_{\hat{x}}} (||\nabla_{\hat{x}} D(\hat{x})||_{2} - 1)^{2} + \lambda_{2} E_{X \sim P_{r}, G(z) \sim P_{G}} \left[ \left( \frac{|D(X) - D(G(z))|}{|D(X - G(z))|} \right) - 1 \right]^{2}$$

### Natural Image Quality Evaluator (NIQE)

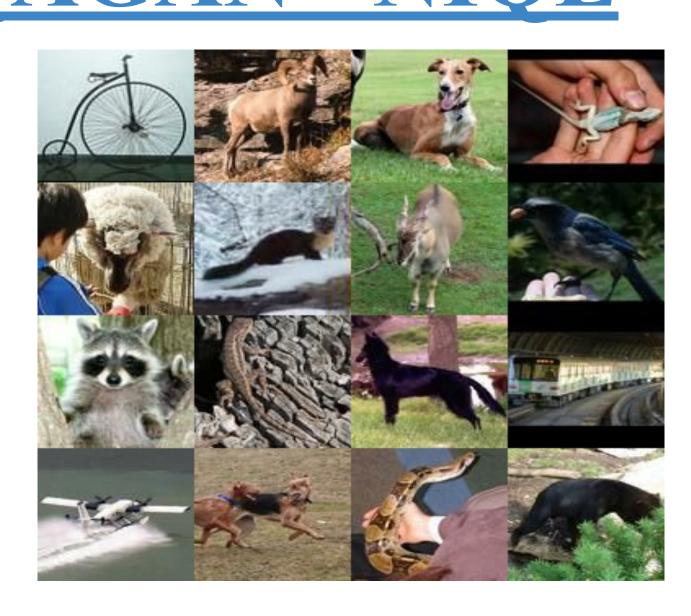
- NIQE builds a statistical model for the class of pristine natural scenes.
- Mean subtracted contrast normalization:

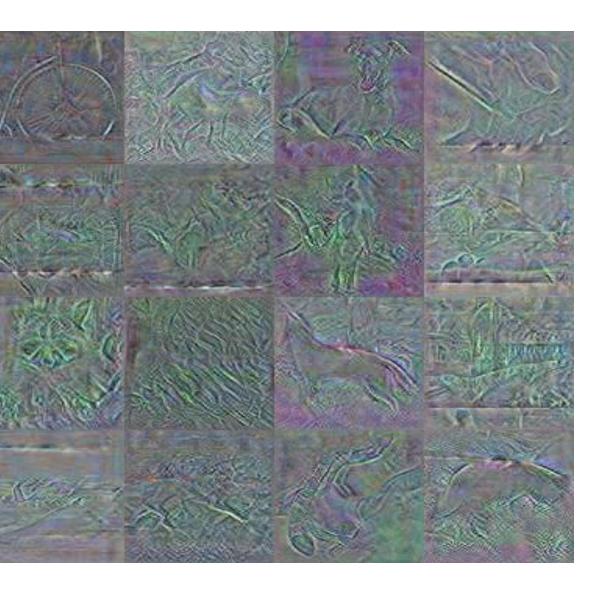
$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}$$

- 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.

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

#### OAGAN - NIOE





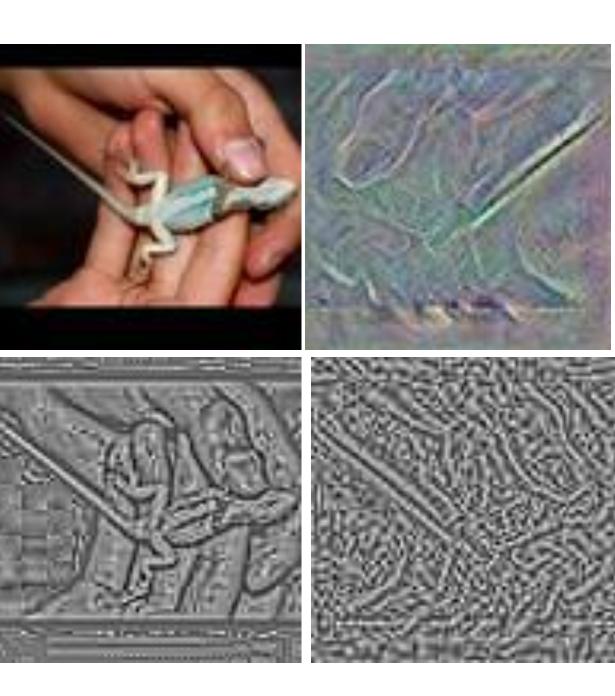
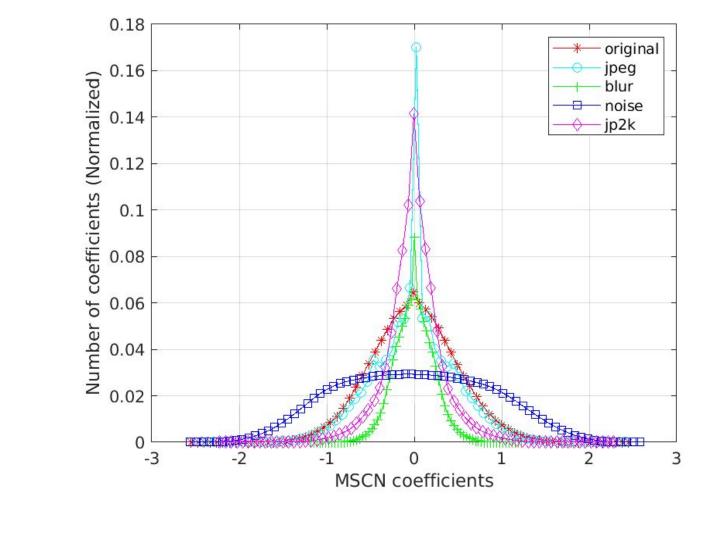
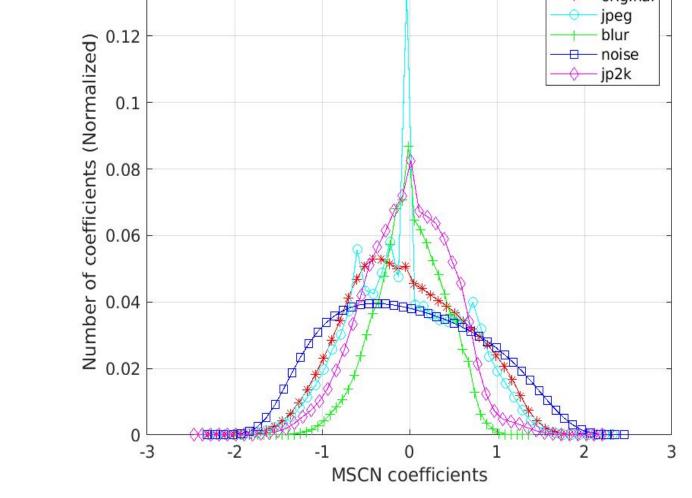


Figure: Real images and corresponding discriminator gradient maps.

Figure: MSCN coeffiecients.





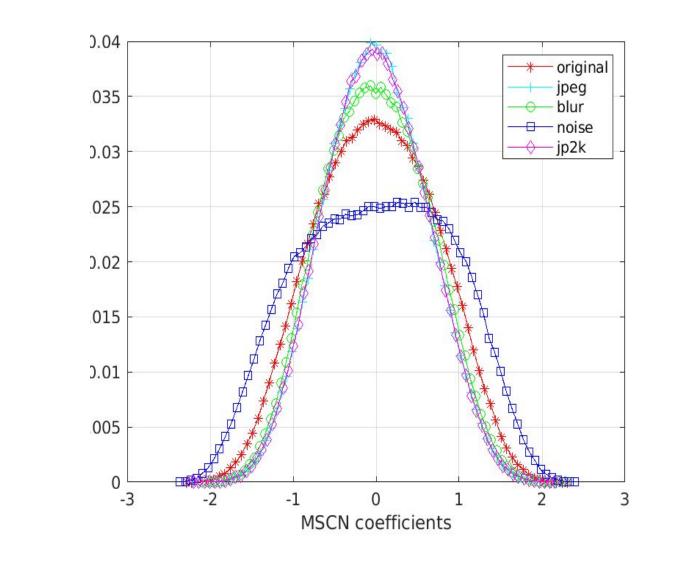


Figure: The empirical histograms of MSCN coefficients

- 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.

$$L_{d} = \min_{D \in \mathcal{D}} \left( E_{z \sim P_{z}} D(G(z)) - E_{X \sim P_{r}} D(X) \right) + \lambda_{1} E_{\hat{x} \sim P_{\hat{x}}} (||\nabla_{\hat{x}} D(\hat{x})||_{2} - 1)^{2} + \lambda_{2} E_{\hat{x} \sim P_{\hat{x}}} (||(\nabla_{\hat{x}} D(\hat{x})|\mu_{P}, \Sigma_{P})||_{\text{NIQE}})$$

 $\lambda_2 E_{X \sim P_r, G(z) \sim P_G} \left[ \left( \frac{|D(X) - D(G(z))|}{d^Q(X, G(z))} \right) - 1 \right]^2$  • where  $\lambda$  '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.

FID		Model	IS	FID	Model	IS	FII
1.09	Rea				Real data	$11.24 \pm 0.12$	7.80
					DCGAN	$6.16 \pm 0.07$	-
	-			3	WGAN-GP	$7.86 \pm 0.10$	$40.2 \pm$
10.5				Programmy Shoral Advances	CTGAN	$8.12 \pm 0.12$	-
6.70	N	MMD GAN-rep	$9.36 \pm 0.0$	$36.67 \pm 0.0$			21.5 ±
	Q.	AGAN (SSIM)	$9.29 \pm 0.05$	$19.77 \pm 0.0091$	The second secon	10-10 (Carlotte Carlotte Carlo	21.0 ⊥
6.421				THE REAL PROPERTY OF THE PERSON OF THE PERSO	$W^{-\frac{1}{2},2}$ - Banach WGAN	$8.26 \pm 0.07$	-
6 504	QF	AGAN (NIQE)	9.1720 ± 0.08	19.45 ± 0.0015	$L^{10}$ - Banach WGAN	$8.31 \pm 0.07$	-
0.004	Į,				MMD GAN-rep-b	$8.29 \pm 0.0$	16.21 ±
					QAGAN (SSIM)	$\textbf{8.37} \pm \textbf{0.04}$	13.91 $\pm$
	FID 1.09 12.89 10.5 6.79 <b>6.421</b> <b>6.504</b>	1.09 Res 12.89 10.5 6.79 QA QA	1.09 Real Data (48 × 48) WGAN-GP SNGAN MMD GAN-rep QAGAN (SSIM) QAGAN (NIQE)	1.09       Real Data $(48 \times 48)$ $26.08 \pm 0.26$ WGAN-GP $9.05 \pm 0.12$ SNGAN $9.10 \pm 0.04$ MMD GAN-rep $9.36 \pm 0.0$ QAGAN (SSIM) $9.29 \pm 0.05$ QAGAN (NIQE) $9.1720 \pm 0.08$	1.09       Real Data $(48 \times 48)$ $26.08 \pm 0.26$ 7.9         WGAN-GP $9.05 \pm 0.12$ $55.1 \pm 0.0$ SNGAN $9.10 \pm 0.04$ $40.10 \pm 0.50$ MMD GAN-rep $9.36 \pm 0.0$ $36.67 \pm 0.0$ QAGAN (SSIM) $9.29 \pm 0.05$ $19.77 \pm 0.0091$ QAGAN (NIQE) $9.1720 \pm 0.08$ $19.45 \pm 0.0013$	$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$

### 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

## Références:

3. Martin Arjovsky, Soumith Chintala, and Lon Bottou. Wasserstein generative adver- sarial networks. In: International Conference on Machine

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).

runet, Edward R Vrscay, and Zhou Wang. On the mathematical properties of the structural similarity index. In: IEEE Transactions

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