CDVAE: Co-embedding Deep Variational Auto Encoder for Conditional Variational Generation Supplementary Materials

Jiajun Lu, Aditya Deshpande, David Forsyth
University of Illinois at Urbana Champaign
{jlu23, ardeshp2, daf}@illinois.edu

1. Architecture Details

Our CDVAE has a different architecture compared to a CVAE. The detailed architecture of our CDVAE is in Table 1. Write l_{in} for the input layer and l_{out} for the output layer, fc stands for a fully connected layer, $mean$ is the mean of the gaussian distribution of the code space, and var is the variance of the gaussian distribution of the code space. $Sample$ is the process of sampling the gaussian distribution with $mean$ and var. l_{out} is sampled from $mean_4$ and var_4. We use L_2 regularization (or weight decay) for the parameters for MDN model. The learning rate is set to 5×10^{-5} and we use the ADAM optimizer. We initially set the reconstruction cost high, LPP embedding guidance cost high, and MDN cost low. We keep this setting and train for 100 epochs. For the next 200 epochs, we gradually decrease the embedding cost, and increase the MDN cost. Finally, we keep the relative cost fixed and train another 200 epochs.

2. DVAE

The difference between DVAE [2] and VAE [1] is multiple layers of gaussian latent variables. DVAE for x_c (same for x_g) consists of L layers of latent variables. To generate a sample from the model, we begin at the top-most layer (L) by drawing from a Gaussian distribution to get $z_{c,L}$.

$$P(z_{c,L}) = \mathcal{N}(z_{c,L}|0, I) \tag{1}$$

The mean and variance for the Gaussian distributions at any lower layer is formed by a non-linear transformation of the sample from above layer.

$$\mu_{c,i} = f_{\mu_{c,i}}(z_{c,i+1}) \tag{2}$$

$$\sigma^2_{c,i} = f_{\sigma^2_{c,i}}(z_{c,i+1}) \tag{3}$$

where f represents multi-layer perceptrons. We descend through the hierarchy by one hot vector sample process.

$$z_{c,i} = \mu_{c,i} + \xi_i \sigma_{c,i} \tag{4}$$

where ξ_i are mutually independent Gaussian variables. x_c is generated by sampling from the Gaussian distribution at the lowest layer.

$$P(x_c|z_{c,1}) = \mathcal{N}(x_c|\mu_{c,0}, \sigma^2_{c,0}) \tag{5}$$

The joint probability distribution $P(x_c, z_c)$ of this model is formulated as

$$\begin{aligned} P(x_c, z_c) &= P(x_c|z_{c,1})P(z_c) \\ &= P(x_c|z_{c,1})P(z_{c,L}) \prod_{i=1}^{L-1} P(z_{c,i}|z_{c,i+1}) \end{aligned} \tag{6}$$

where $P(z_{c,i}|z_{c,i+1}) = \mathcal{N}(z_{c,i}|\mu_{c,i}, \sigma^2_{c,i})$. Other details of the DVAE model are similar to VAE.

2.1. Inference

DVAE with several layers of dependent stochastic variables are difficult to train which limits the improvements obtained using these highly expressive models. LVAE [3] recursively corrects the generative distribution by a data dependent approximate likelihood in a process resembling the recent Ladder Network. It utilizes a deeper more distributed hierarchy of latent variables and captures more complex structures. We follow this work and for x_c, write $\mu_{c,p,i}$ and $\sigma^2_{c,p,i}$ for the mean and variance on the i's level of generative side, write $\mu_{c,q,i}$ and $\sigma^2_{c,q,i}$ for the mean and variance on the i's level of inference side.

This changes the notation in the previous part on the generative side.

$$P_p(z_c) = P_p(z_{c,L}) \prod_{i=1}^{L-1} P_p(z_{c,i}|z_{c,i+1}) \tag{7}$$

$$P_p(z_{c,L}) = \mathcal{N}(z_{c,L}|0, I) \tag{8}$$

$$P_p(z_{c,i}|z_{c,i+1}) = \mathcal{N}(z_{c,i}|\mu_{c,p,i}, \sigma^2_{c,p,i}) \tag{9}$$

Layers	Conditional DVAE		MDN$_x$ (x is GMM num)	Generative DVAE	
l_{in}	(None, 1024)			(None, 1024)	
$fc1$	(1024, 512)			(1024, 512)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc2$	(512, 512)			(512, 512)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc3_1, fc3_2$	(512, 64)	(512, 64)		(512, 64)	(512, 64)
$activation$	Identity	SoftPlus		Identity	SoftPlus
$mean_1, var_1$	(None, 64)	(None, 64)		(None, 64)	(None, 64)
$sample_1$	(None, 64)			(None, 64)	
$fc4$	(64, 256)			(64, 256)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc5$	(256, 256)			(256, 256)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc6_1, fc6_2$	(256, 32)	(256, 32)		(256, 32)	(256, 32)
$activation$	Identity	SoftPlus		Identity	SoftPlus
$mean_2, var_2$	(None, 32)	(None, 32)		(None, 32)	(None, 32)
$sample_2$	(None, 32)		$fc_a(32, (32+1)x)$ = $fc_a(32, (dim(\mu_k) + dim(\pi_k))x)$, activation=tanh, $fc_b((32+1)x, (32+1)x)$, activation = tanh, cost = GMM($z_g \mid \mu_k, \pi_k, x$)	(None, 32)	
$fc7$	(32, 256)			(32, 256)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc8$	(256, 256)			(256, 256)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc9_1, fc9_2$	(256, 64)	(256, 64)		(256, 64)	(256, 64)
$activation$	Identity	SoftPlus		Identity	SoftPlus
$mean_3, var_3$	(None, 64)	(None, 64)		(None, 64)	(None, 64)
$sample_3$	(None, 64)			(None, 64)	
$fc10$	(64, 512)			(64, 512)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc11$	(512, 512)			(512, 512)	
$activation$	Leaky Rectify			Leaky Rectify	
$fc12_1, fc12_2$	(512,1024)	(512,1024)		(512,1024)	(512,1024)
$activation$	Identity	SoftPlus		Identity	SoftPlus
$mean_4, var_4$	(None, 1024)	(None, 1024)		(None, 1024)	(None, 1024)
l_{out}	(None, 1024)			(None, 1024)	

Table 1. Details for the CDVAE architecture we proposed.

$$P_p(x_c|z_{c,1}) = \mathcal{N}(x_c|\mu_{c,p,0}, \sigma^2_{c,p,0}) \tag{10}$$

On the inference side, the notation also changes.

$$P_q(z_c|x_c) = P_q(z_{c,1}|x_c)\prod_{i=2}^{L} P_q(z_{c,i}|z_{c,i-1}) \tag{11}$$

$$P_q(z_{c,1}|x_c) = \mathcal{N}(z_{c,1}|\mu_{c,q,1}, \sigma^2_{c,q,1}) \tag{12}$$

$$P_q(z_{c,i}|z_{c,i-1}) = \mathcal{N}(z_{c,i}|\mu_{c,q,i}, \sigma^2_{c,q,i}) \tag{13}$$

During inference, first a deterministic upward pass computes the approximate distribution $\hat{\mu}_{c,q,i}$ and $\hat{\sigma}^2_{c,q,i}$. This is followed by a stochastic downward pass recursively computing both the approximate posterior and generative distributions.

$$P_q(z_c|x_c) = P_q(z_{c,L}|x_c)\prod_{i=1}^{L-1} P_q(z_{c,i}|z_{c,i+1}) \tag{14}$$

$$\sigma_{c,q,i} = \frac{1}{\hat{\sigma}^{-2}_{c,q,i} + \sigma^{-2}_{c,p,i}} \tag{15}$$

$$\mu_{c,q,i} = \frac{\hat{\mu}_{c,q,i}\hat{\sigma}^{-2}_{c,q,i} + \mu_{c,p,i}\sigma^{-2}_{c,p,i}}{\hat{\sigma}^{-2}_{c,q,i} + \sigma^{-2}_{c,p,i}} \tag{16}$$

$$P_q(z_{c,i}|\cdot) = \mathcal{N}(z_{c,i}|\mu_{c,q,i}, \sigma^2_{c,q,i}) \tag{17}$$

where $\mu_{c,q,L} = \hat{\mu}_{c,q,L}$ and $\sigma^2_{c,q,L} = \hat{\sigma}^2_{c,q,L}$.

3. Joint Models

First, we prove that if the joint probability is independent, we will get two separate DVAEs. Then, we prove the derivations for joint model with non-independent joint probability.

3.1. Separate DVAEs

From Section 3 in the paper, the joint probability $P(x_c, x_g)$ in CDVAE model is

$$P(x_c, x_g) = \int_z P(x_c|z_c)P(x_g|z_g)P(z_g, z_c)dz_g dz_c \tag{18}$$

If z_c and z_g are independent, so $P(z_g, z_c) = P(z_g)P(z_c)$, and Equation 18 can be transformed

$$\begin{aligned} P(x_c, x_g) &= \int_z P(x_c|z_c)P(x_g|z_g)P(z_g)P(z_c)dz_g dz_c \\ &= \int_z (P(x_c|z_c)P(z_c))dz_c(P(x_g|z_g)P(z_g))dz_g \\ &= \int_{z_c} P(x_c|z_c)P(z_c)dz_c + \int_{z_g} P(x_g|z_g)P(z_g)dz_g \\ &= P(x_c) + P(x_g) \end{aligned} \tag{19}$$

where $P(x_c)$ is DVAE model for x_c and $P(x_g)$ is DVAE model for x_g.

3.2. Joint Model Derivation

From Section 2 in the paper, we have objective function for VAE as

$$\text{VAE}(\theta) = \sum_{data}[\mathbb{E}_Q \log P(x|z) - \mathbb{D}(Q||P(z))] \tag{20}$$

where $Q = P(z|x)$. Applying the same derivations, the objective function for our CDVAE model can be written as

$$\text{CDVAE}(\theta_c, \theta_g) = \textstyle\sum_{data}[\mathbb{E}_Q \log P(x_c, x_g|z_c, z_g) - \mathbb{D}(Q||P(z_c, z_g))] \tag{21}$$

where $Q = P(z_c, z_g|x_c, x_g)$. Assume it is possible to encode x_c without seeing x_g, then the variational distribution $Q = P(z_c|x_c)P(z_g|x_g)$ applies. It is also possible to decode x_c without seeing x_g, so we have $P(x_c, x_g|z_c, z_g) = P(x_c|z_c)P(x_g|z_g)$. With these formulas, Equation 21 can be transformed

$$\begin{aligned} \mathbb{E}_Q \log P(x_c, x_g|z_c, z_g) &= \mathbb{E}_Q \log(P(x_c|z_c)P(x_g|z_g)) \\ &= \mathbb{E}_{Q_1} \log P(x_c|z_c) \\ &+ \mathbb{E}_{Q_2} \log P(x_g|z_g) \end{aligned} \tag{22}$$

where $Q_1 = P(z_c|x_c)$ and $Q_2 = P(z_g|x_g)$. The joint distribution can be written as $\log P(z_c, z_g) = \log P(z_c) + \log P(z_g) + F_{\text{mdn}}(z_c, z_g)$, so we have the following equations for the second part.

$$\begin{aligned} \mathbb{D}(Q||P(z_c, z_g)) &= \mathbb{D}(P(z_c|x_c)P(z_g|x_g)||P(z_c, z_g)) \\ &= \mathbb{E}_Q(\log(P(z_c|x_c)P(z_g|x_g)) - \log P(z_c, z_g)) \\ &= \mathbb{E}_{Q_1}(\log P(z_c|x_c)) + \mathbb{E}_{Q_2}(\log P(z_g|x_g)) \\ &- \mathbb{E}_{Q_1}(\log P(z_c)) - \mathbb{E}_{Q_2}(\log P(z_g)) \\ &- \mathbb{E}_Q(F_{\text{mdn}}(z_c, z_g)) \\ &= \mathbb{D}(Q_1||P(z_c)) + \mathbb{D}(Q_2||P(z_g)) \\ &- \mathbb{E}_Q(F_{\text{mdn}}(z_c, z_g)) \end{aligned} \tag{23}$$

In our CDVAE model, we have $\log P(z_c) = -\frac{z_c^T z_c}{2}$ and $\log P(z_g) = -\frac{z_g^T z_g}{2}$ because z_c and z_g are Gaussian distributions. Our CDVAE objective function turns into

$$\text{CDVAE}(\theta_c, \theta_g) = \text{DVAE}(\theta_c) + \text{DVAE}(\theta_g) + \textstyle\sum_{data} \mathbb{E}_Q(F_{\text{mdn}}(z_c, z_g)) \tag{24}$$

4. Embedding Influence

We compare the results with embedding guidance and without embedding guidance. The comparisons for re-shading can be found in Figures 1, 2, 3 and 4. The re-shading results without embedding guidance tend to have less variety, more flaws and artifacts. The comparisons for re-saturation can be found in Figures 5, 6, 7 and 8. The re-saturation results without embedding guidance tend to have limited variety and produce less vivid results.

5. Quantitative Results

The detailed quantitative evaluation results for **photo relighting** are in Table 2 and **image resaturation** are in Table 3. The tables contain best error to ground-truth with different sample numbers. As the sample number increases, the error drops fast at beginning, and then becomes stable. Our CDVAEs are consistently better than other methods. The second parts of both tables are average variances across 100 samples. We only report the final variance, since it almost does not change with the sample number. The variance we report comes from 100 samples.

	Best Error to Ground Truth					Variance
	Sample# 3	Sample# 10	Sample# 30	Sample# 60	Sample# 100	Sample#
NN	3.04	2.30	1.93	1.76	1.66	1.61
CVAE	2.07	1.83	1.68	1.60	1.56	0.19
CGAN	3.07	2.49	2.16	2.02	1.94	1.19
CPixel	3.06	2.32	1.91	1.74	1.59	1.92
CDVAE$_{noemb}$	2.78	2.19	1.82	1.66	1.57	1.39
CDVAE4	2.44	1.66	1.33	1.20	1.11	1.77
CDVAE12	2.49	1.69	1.33	1.20	1.12	1.74

Table 2. Photo relighting results. First part is best error to ground truth with different sample numbers; second part is variance, which is stable with different sample numbers. (all results need $\times 10^{-2}$)

	Best Error to Ground Truth					Variance
	Sample# 3	Sample# 10	Sample# 30	Sample# 60	Sample# 100	Sample#
NN	10.12	8.40	7.09	6.52	6.20	4.58
CVAE	6.73	5.59	4.93	4.53	4.25	1.20
CGAN	8.06	6.20	5.37	5.03	4.83	4.79
CPixel	7.94	6.43	5.94	5.51	5.29	4.34
CDVAE$_{noemb}$	7.08	5.74	4.95	4.57	4.32	1.53
MDN4	6.62	5.11	4.37	4.05	3.86	3.48
MDN12	6.40	5.04	4.33	4.02	3.82	3.55

Table 3. Image re-saturation results. First part is best error to ground truth with different sample numbers; second part is variance, which is stable with different sample numbers. (all results need $\times 10^{-2}$)

6. Qualitative Results

We include more qualitative results and comparisons in this section. **Photo relighting** results and comparisons can be found in Figure 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 and 25. Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety. **Image re-saturation** results and comparisons can be found in Figure 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36 37, 38, 39, 40, 41 and 42. Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

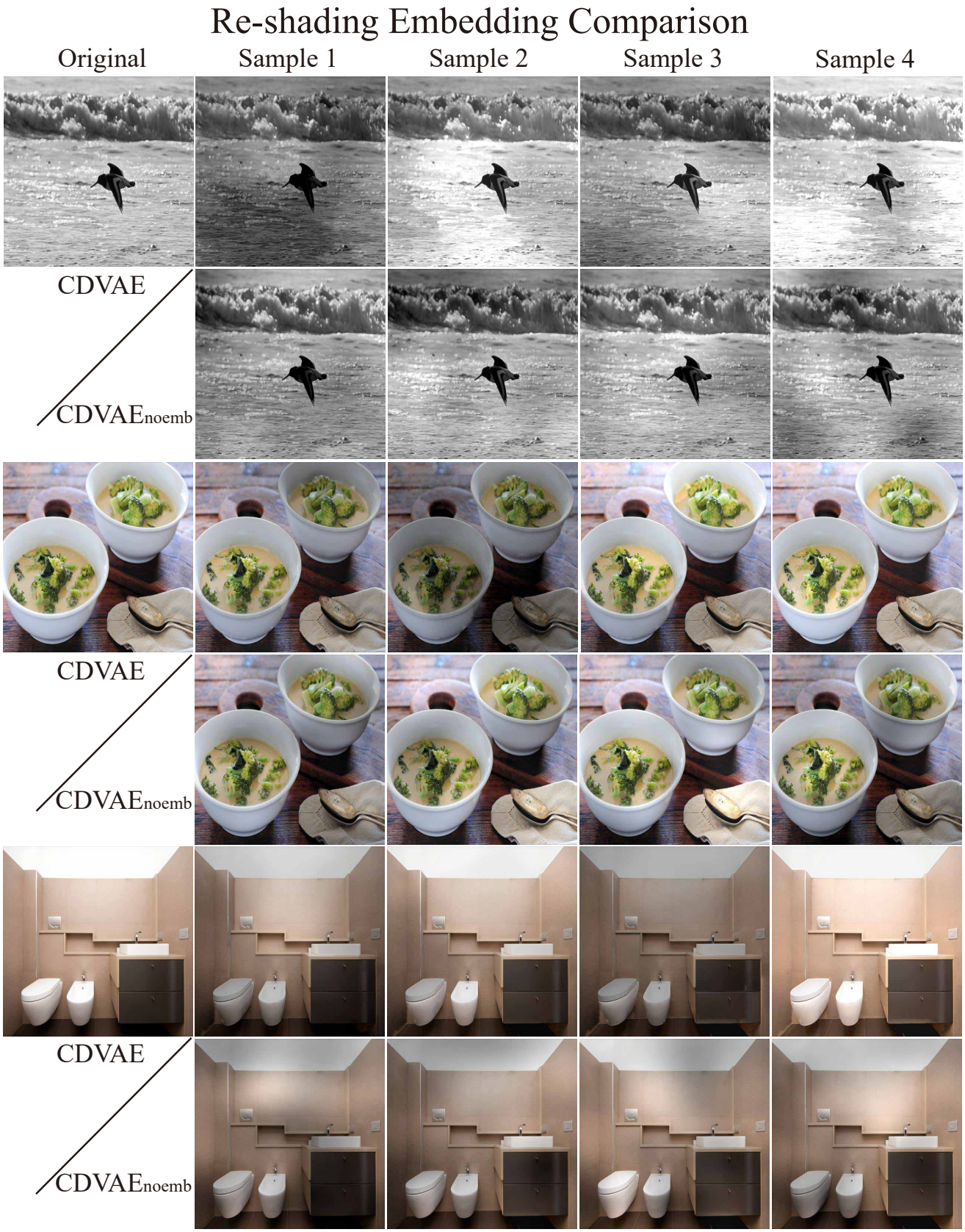

Figure 1. Comparisons to no embedding guidance for re-shading results (part 1). The re-shading results without embedding guidance tend to have less variety, more flaws and artifacts.

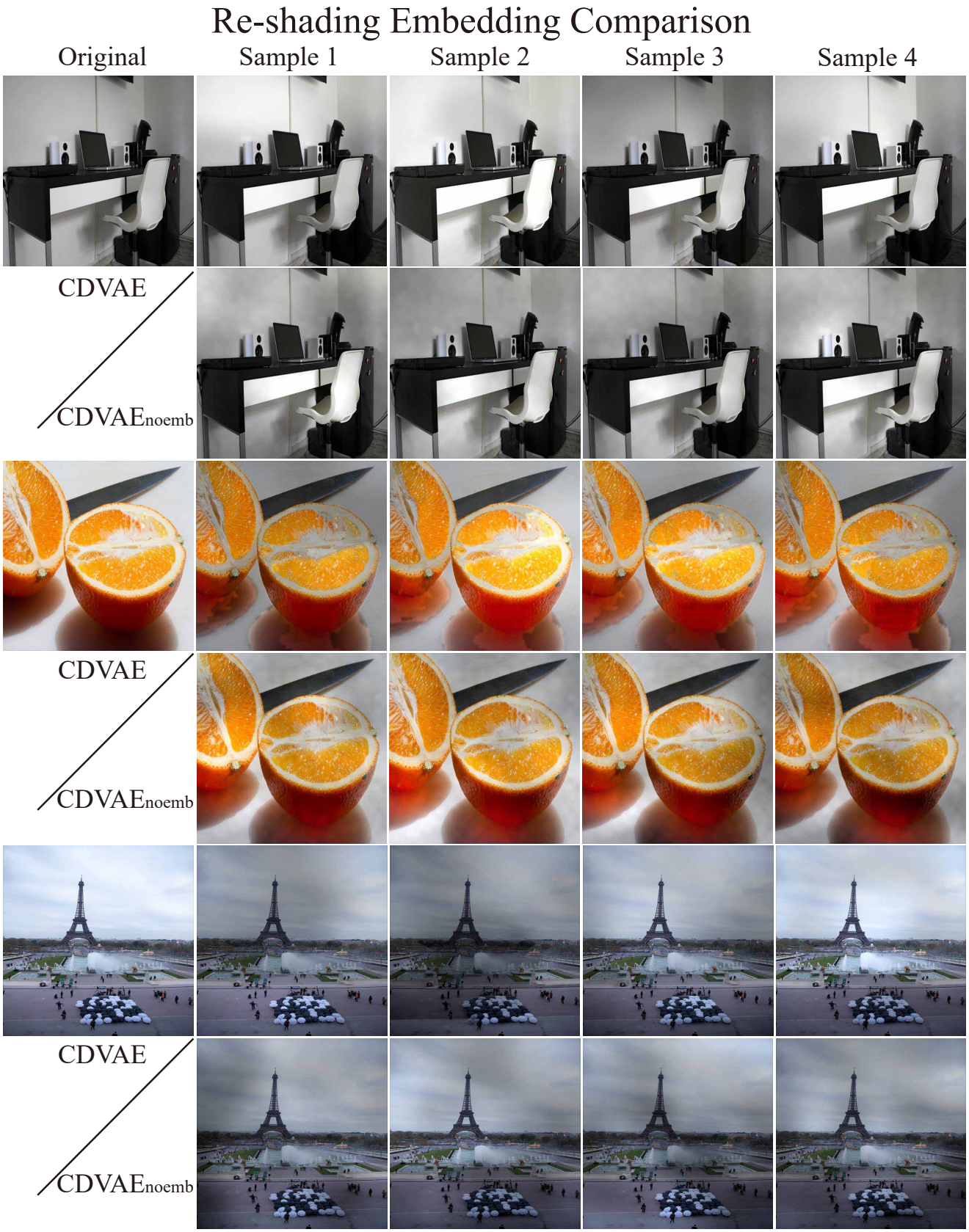

Figure 2. Comparisons to no embedding guidance for re-shading results (part 2). The re-shading results without embedding guidance tend to have less variety, more flaws and artifacts.

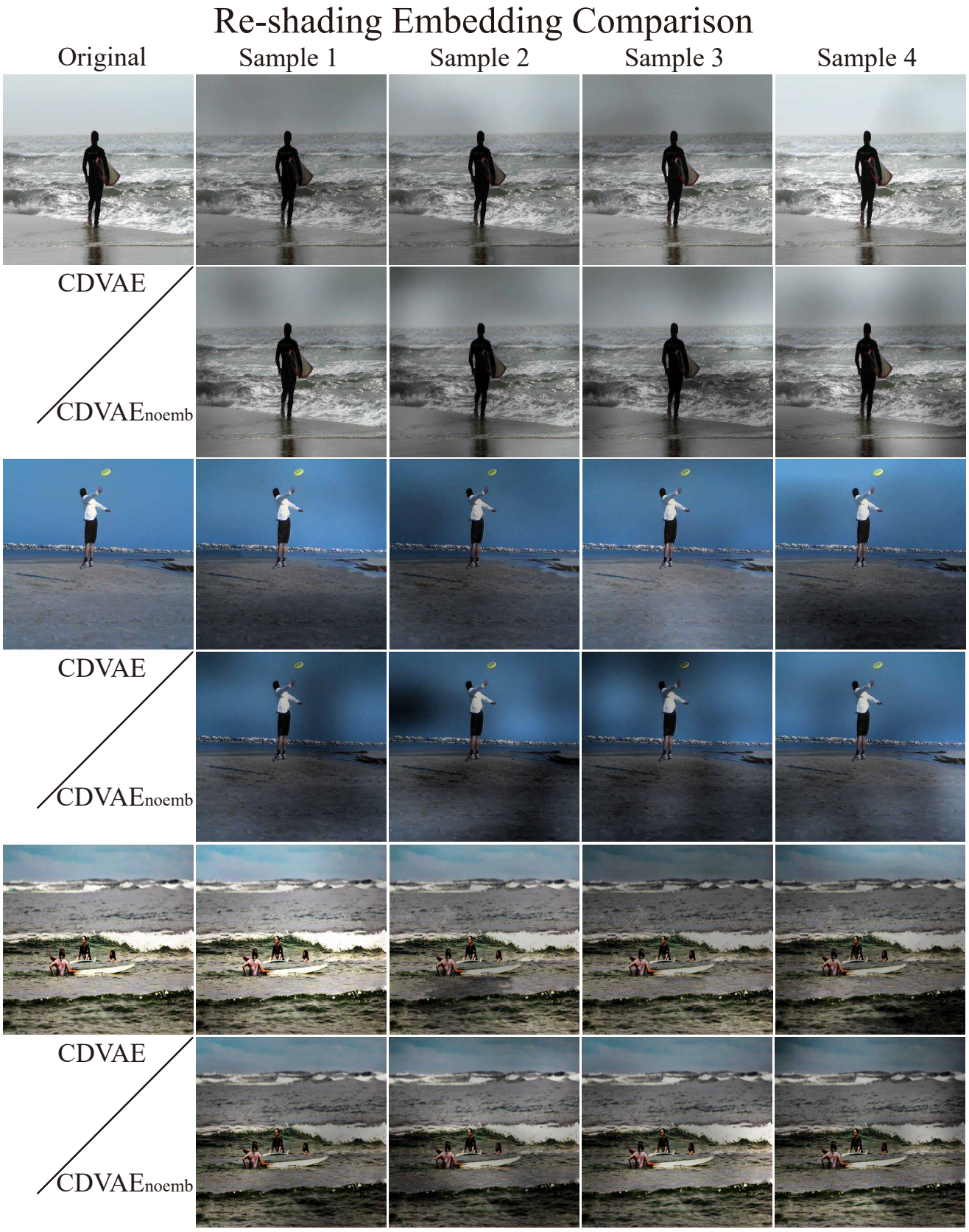

Figure 3. Comparisons to no embedding guidance for re-shading results (part 3). The re-shading results without embedding guidance tend to have less variety, more flaws and artifacts.

Figure 4. Comparisons to no embedding guidance for re-shading results (part 4). The re-shading results without embedding guidance tend to have less variety, more flaws and artifacts.

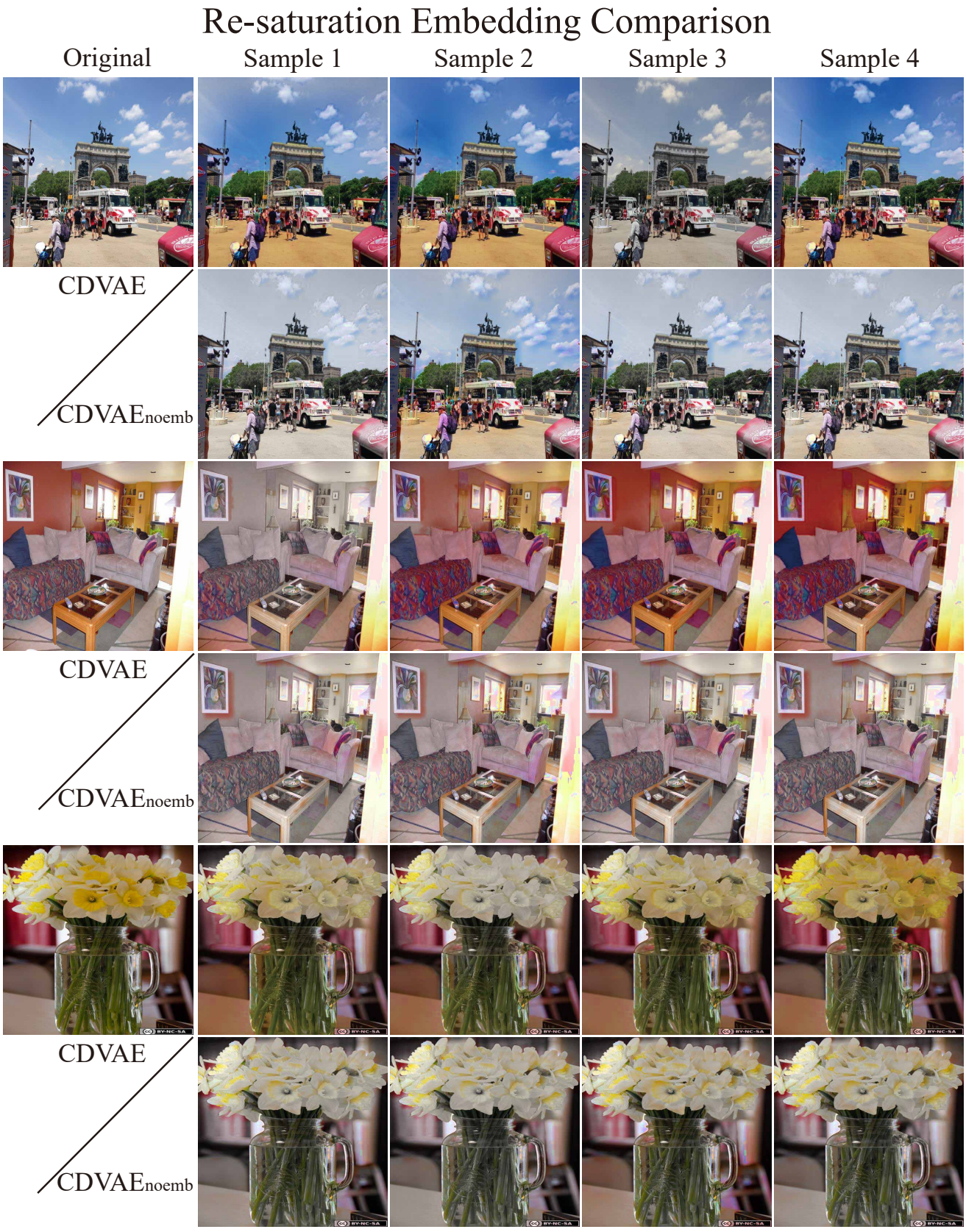

Figure 5. Comparisons to no embedding guidance for re-saturation results (part 1). The re-saturation results without embedding guidance tend to have limited variety and produce less vivid results.

Figure 6. Comparisons to no embedding guidance for re-saturation results (part 2). The re-saturation results without embedding guidance tend to have limited variety and produce less vivid results.

Figure 7. Comparisons to no embedding guidance for re-saturation results (part 3). The re-saturation results without embedding guidance tend to have limited variety and produce less vivid results.

Figure 8. Comparisons to no embedding guidance for re-saturation results (part 4). The re-saturation results without embedding guidance tend to have limited variety and produce less vivid results.

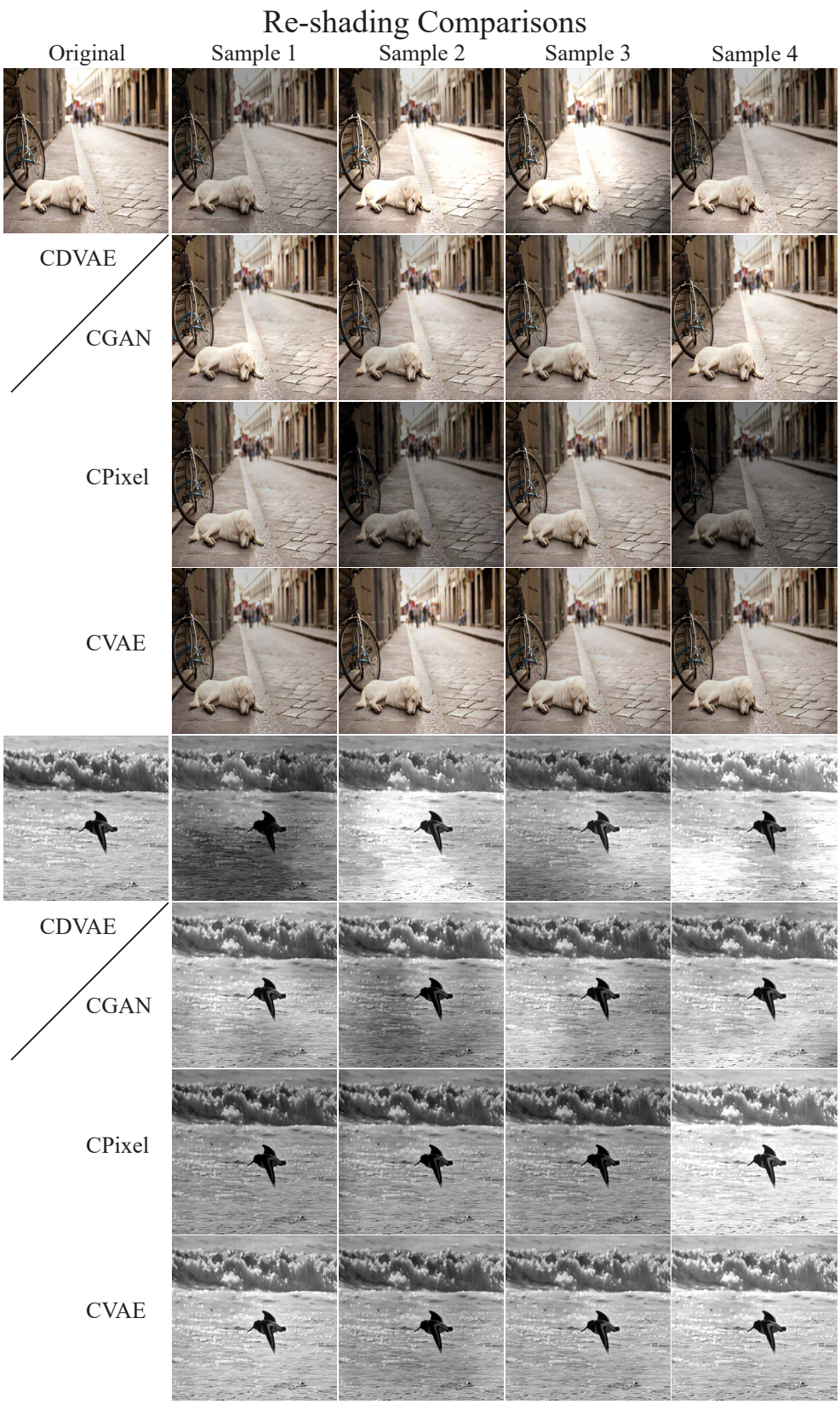

Figure 9. Photo relighting results (part 1). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

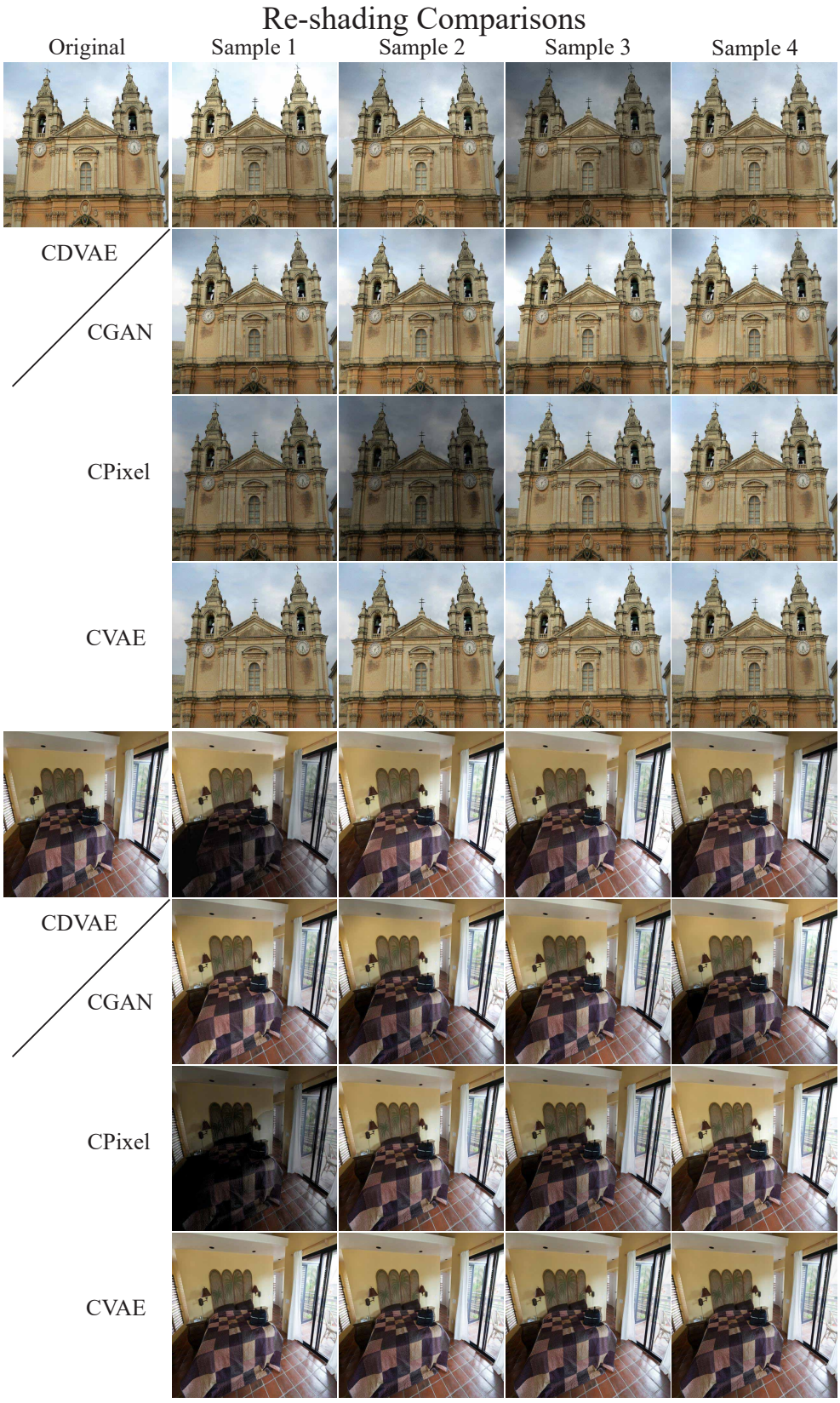

Figure 10. Photo relighting results (part 2). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

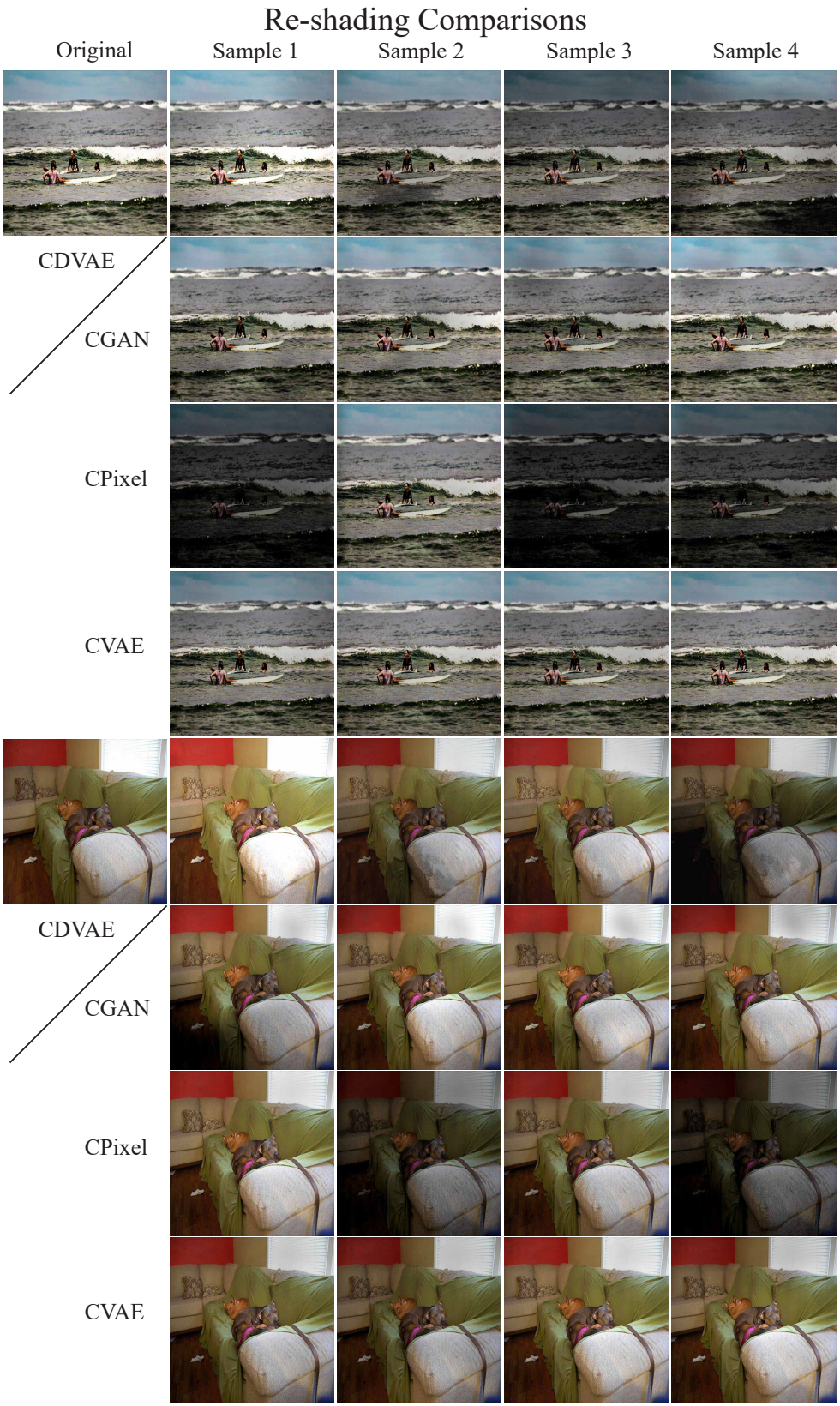

Figure 11. Photo relighting results (part 3). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

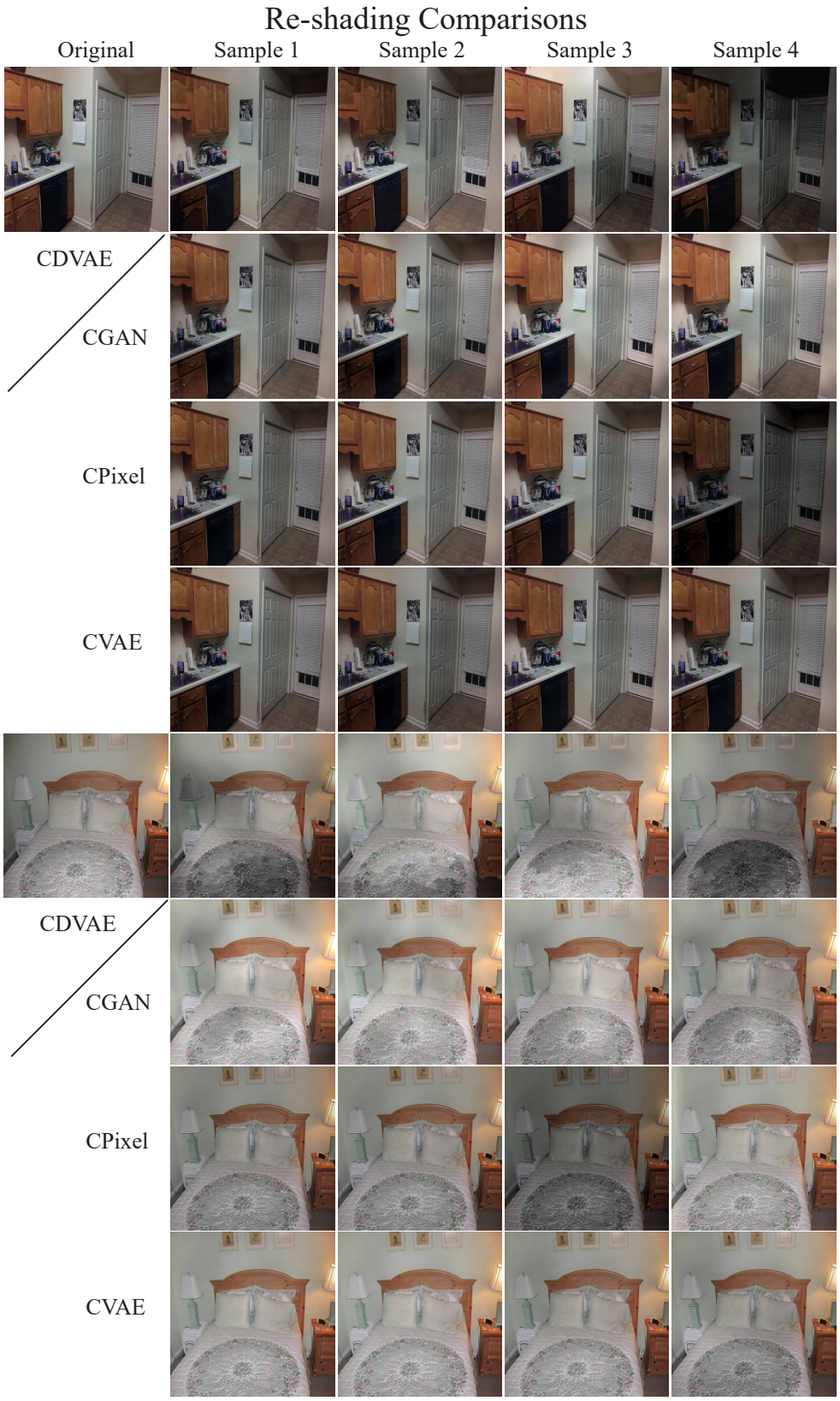

Figure 12. Photo relighting results (part 4). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

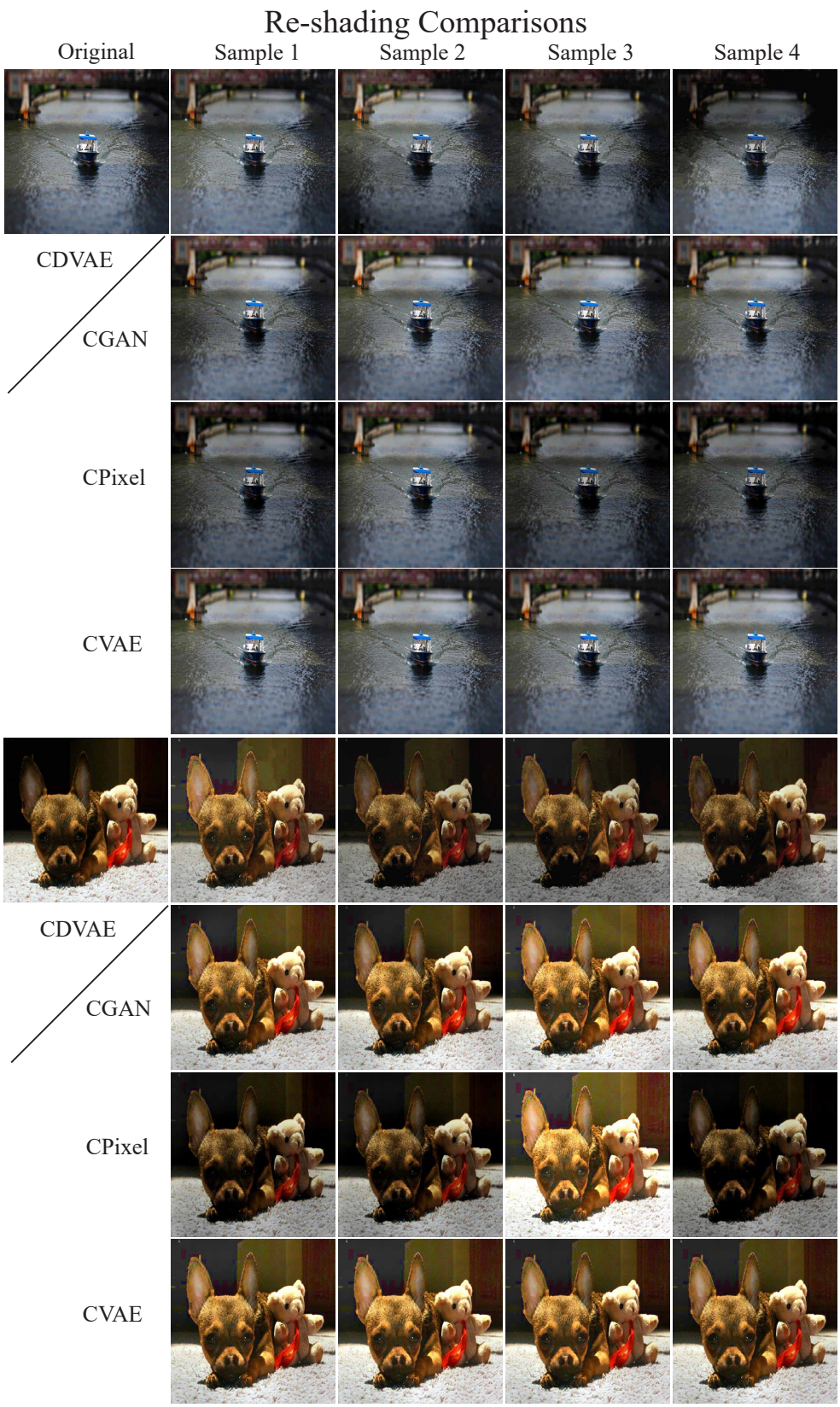

Figure 13. Photo relighting results (part 5). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

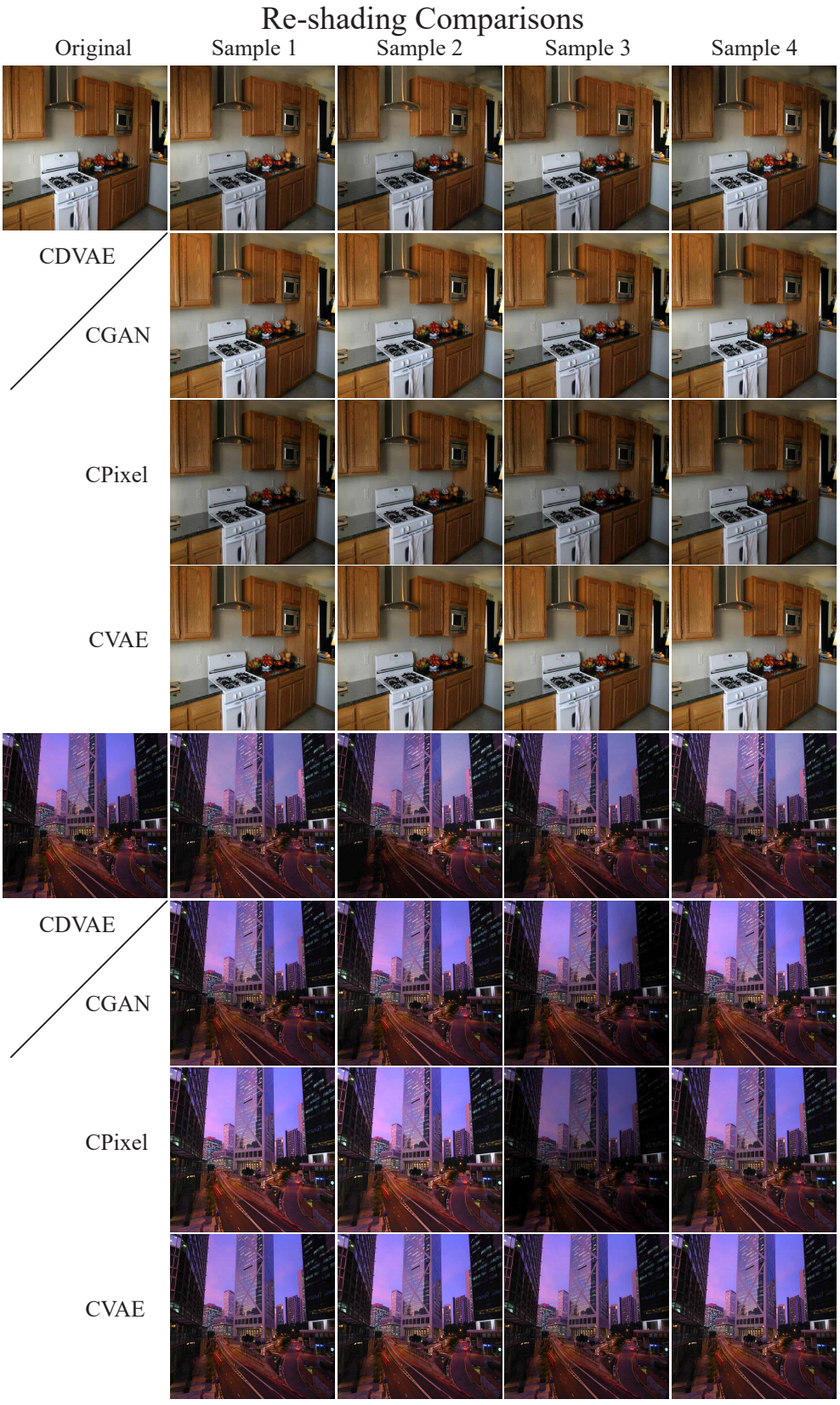

Figure 14. Photo relighting results (part 6). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

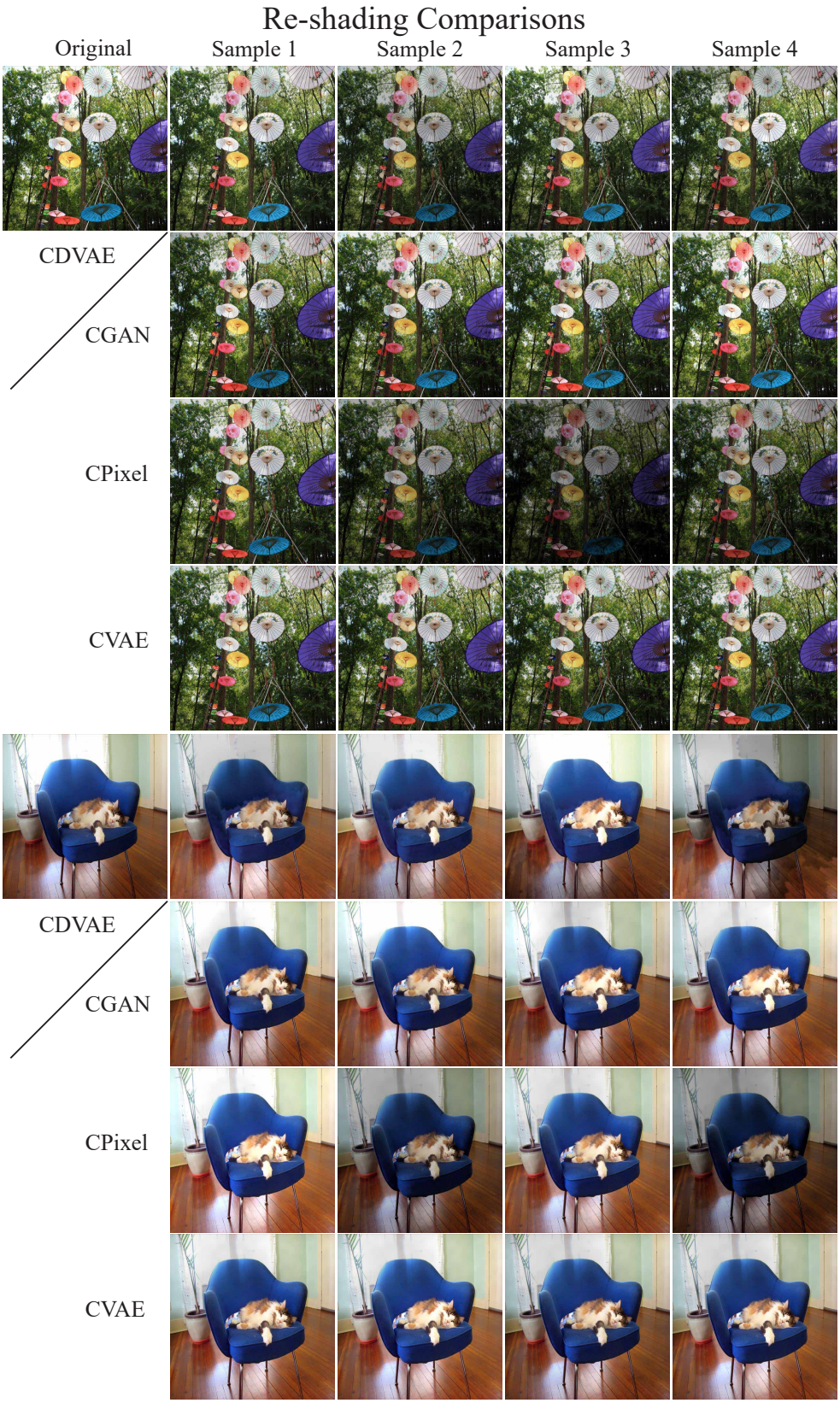

Figure 15. Photo relighting results (part 7). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

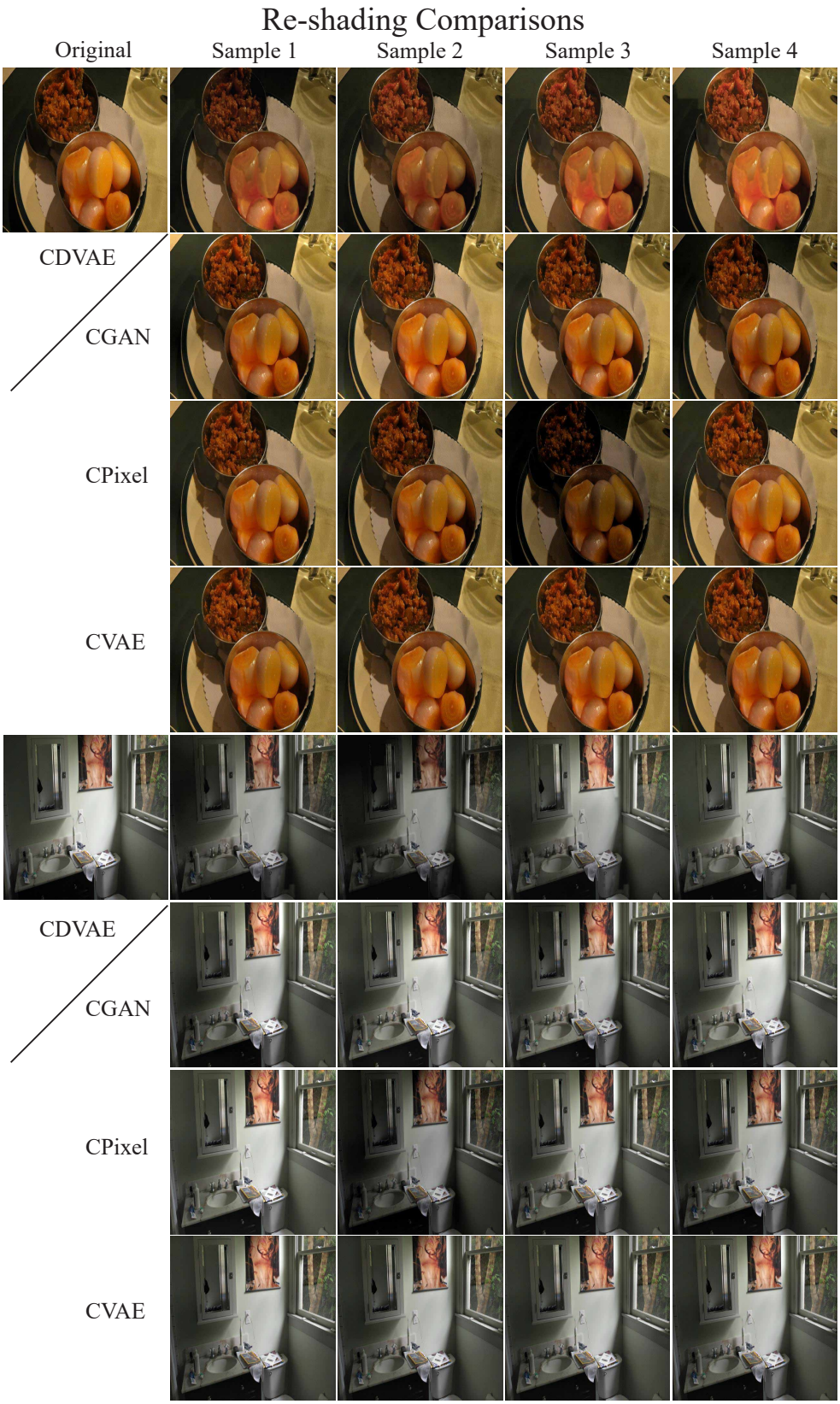

Figure 16. Photo relighting results (part 8). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

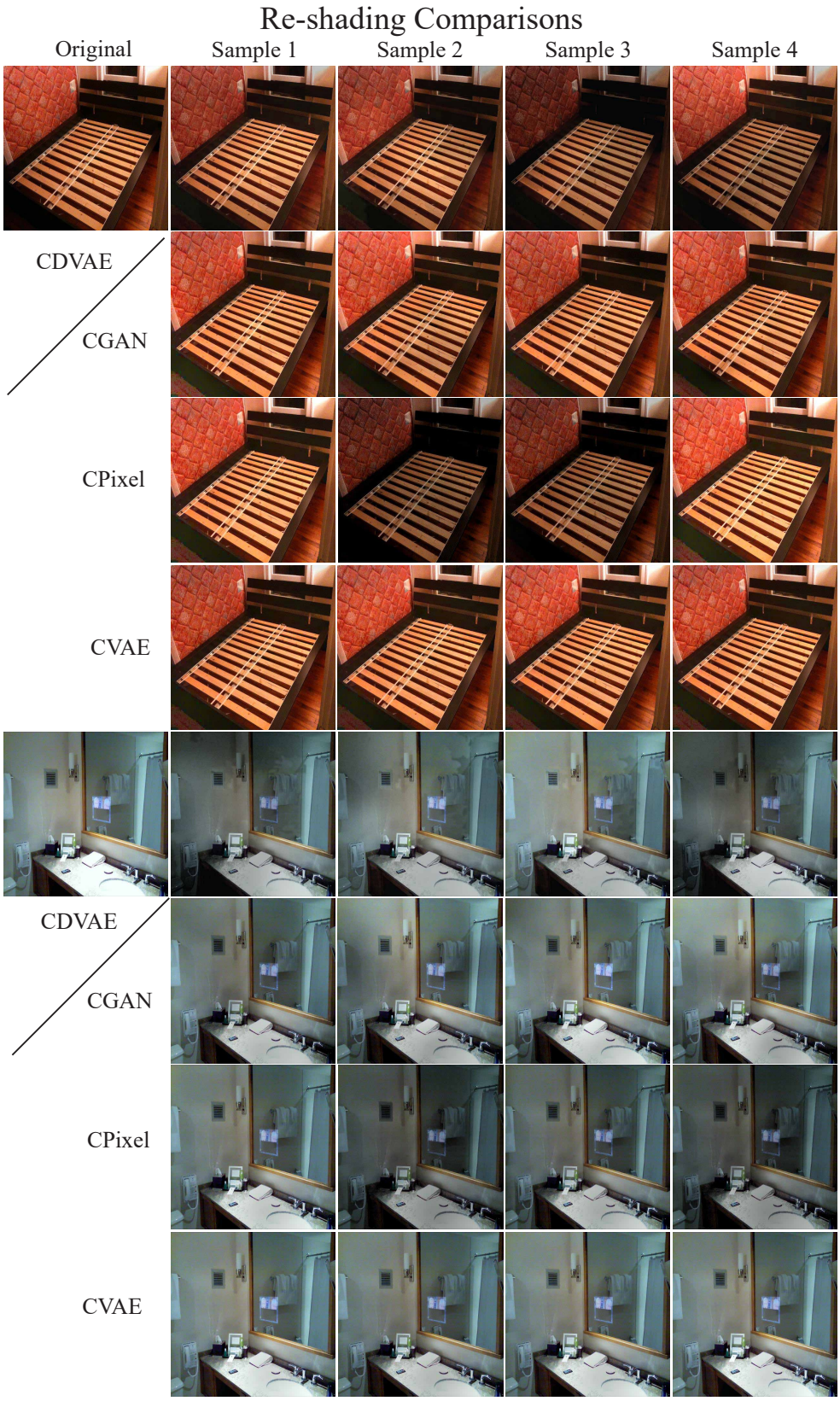

Figure 17. Photo relighting results (part 9). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

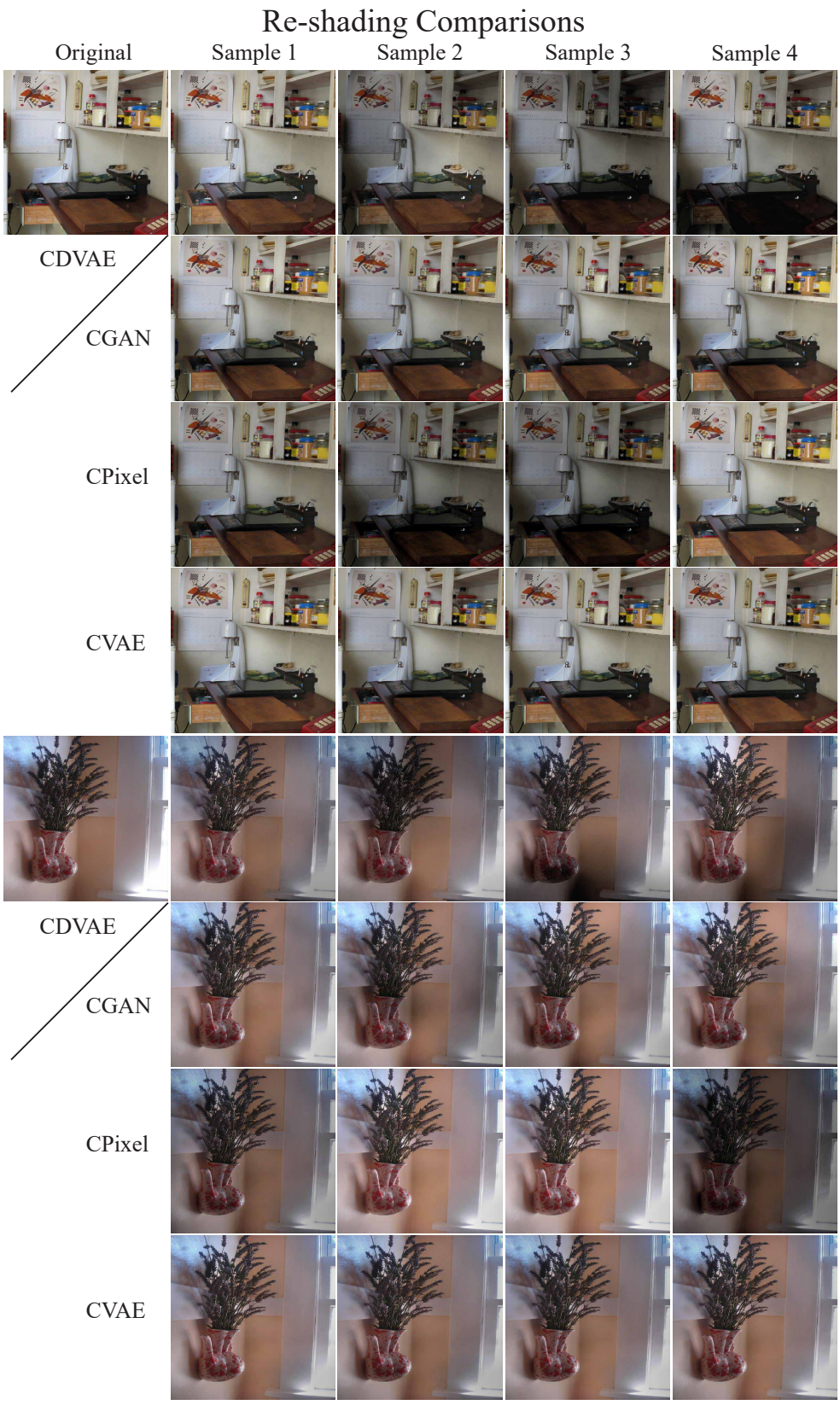

Figure 18. Photo relighting results (part 10). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

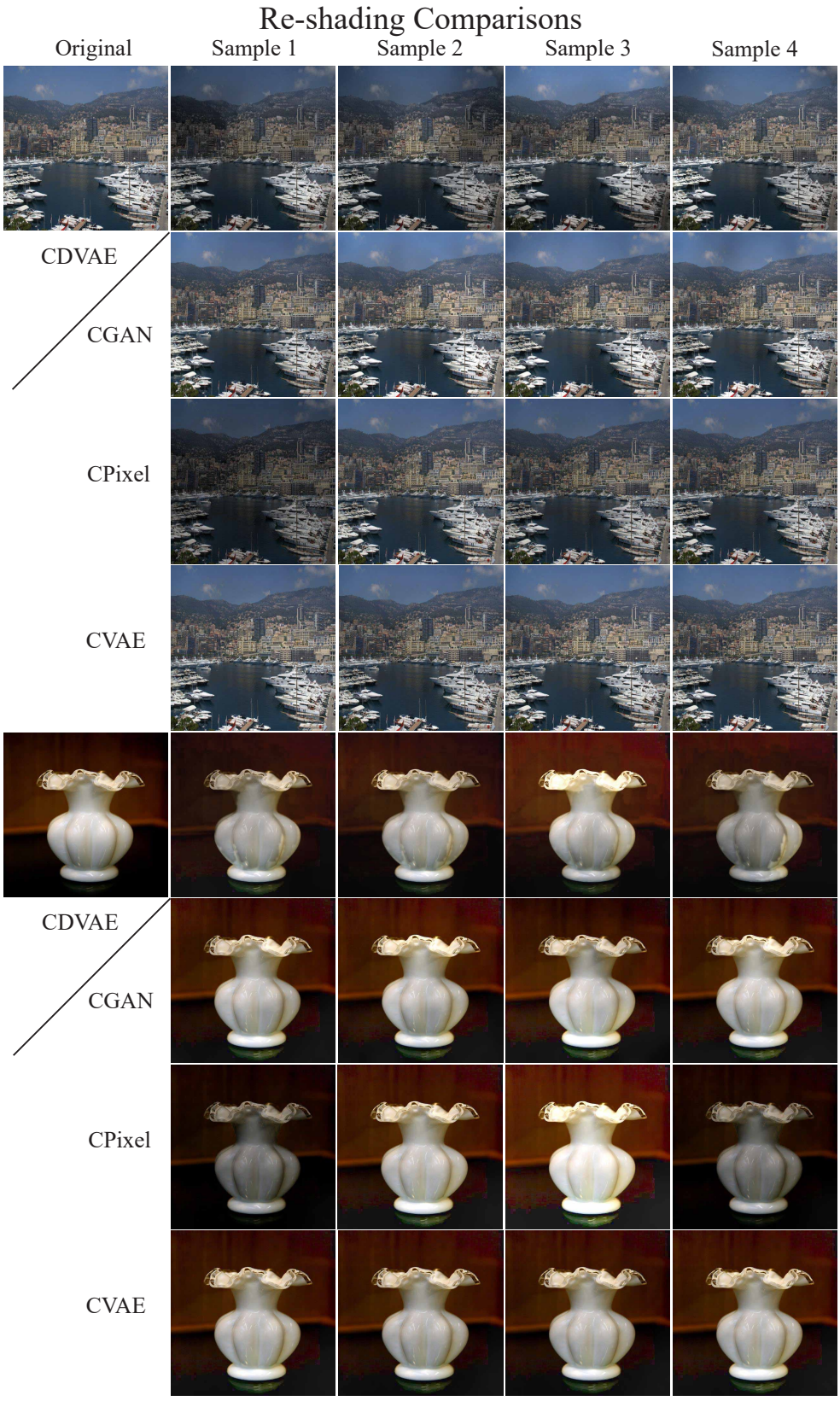

Figure 19. Photo relighting results (part 11). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

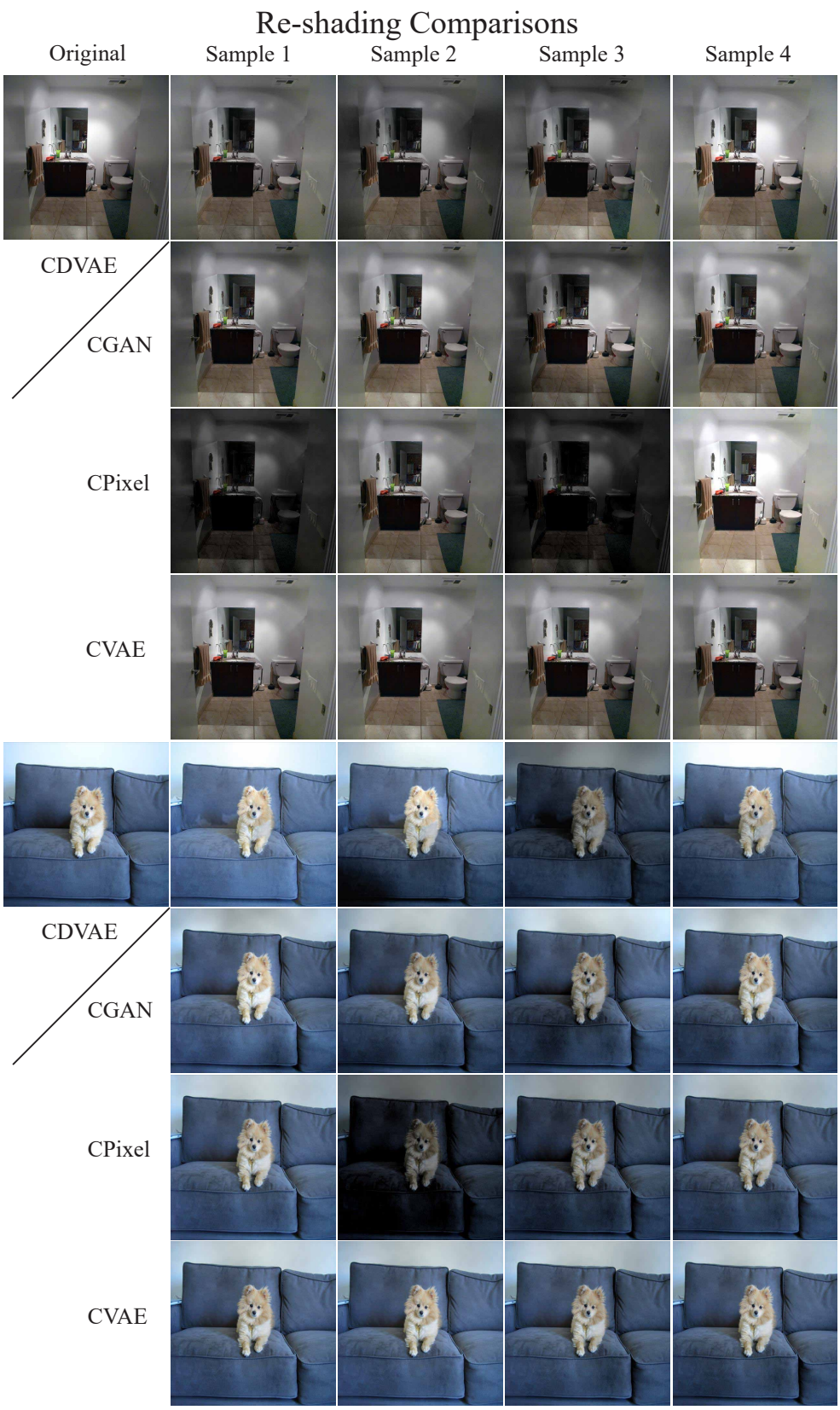

Figure 20. Photo relighting results (part 12). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

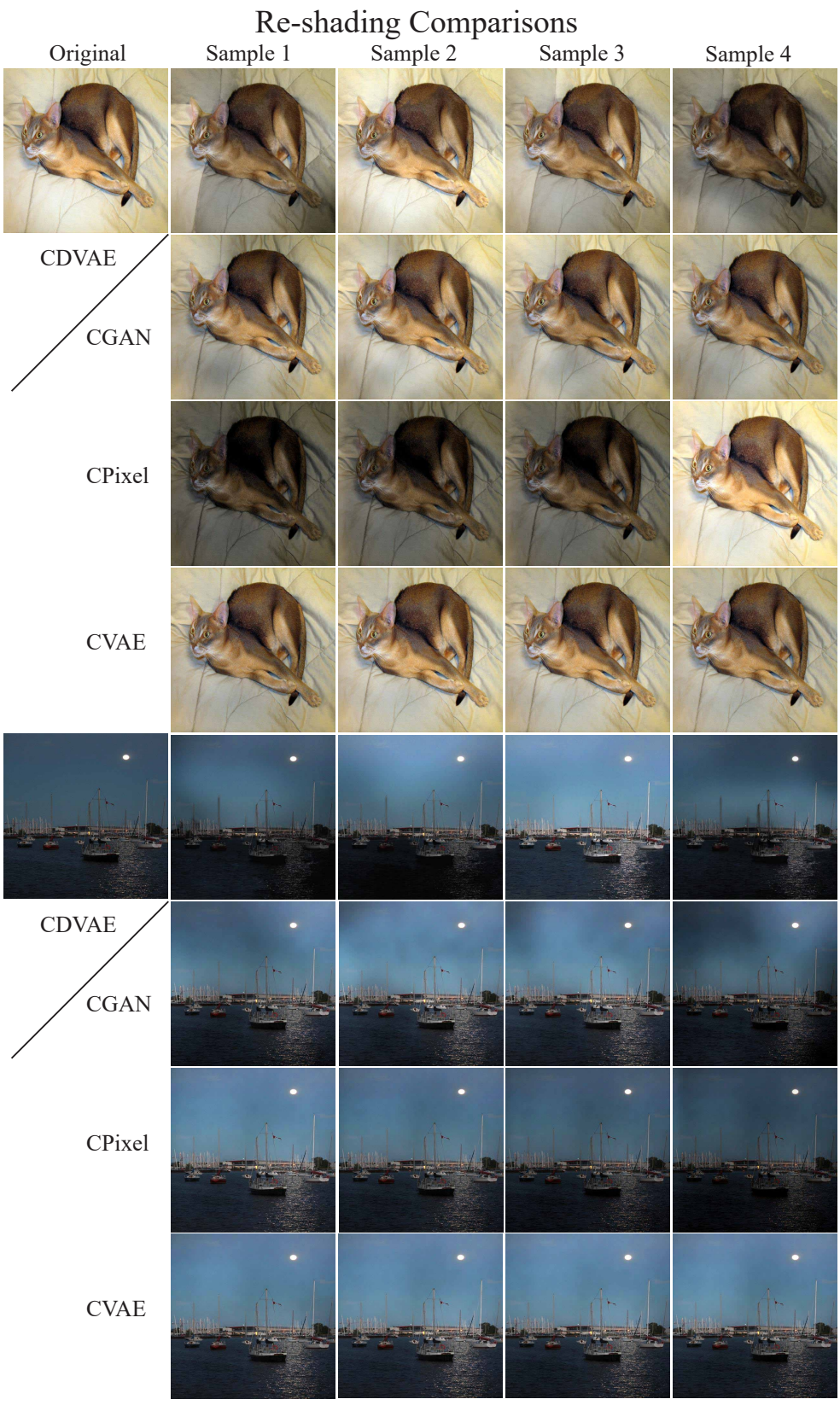

Figure 21. Photo relighting results (part 13). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

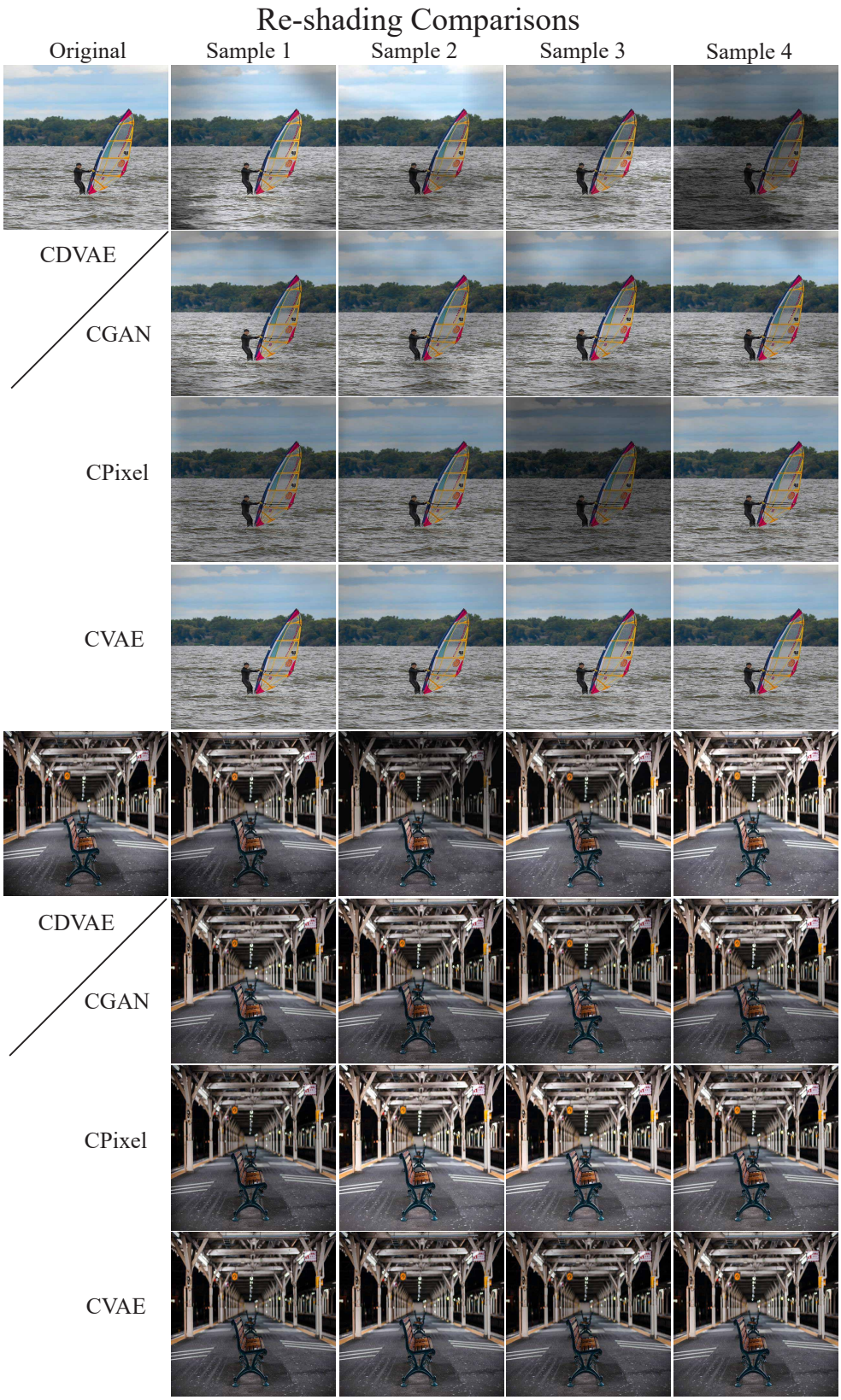

Figure 22. Photo relighting results (part 14). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

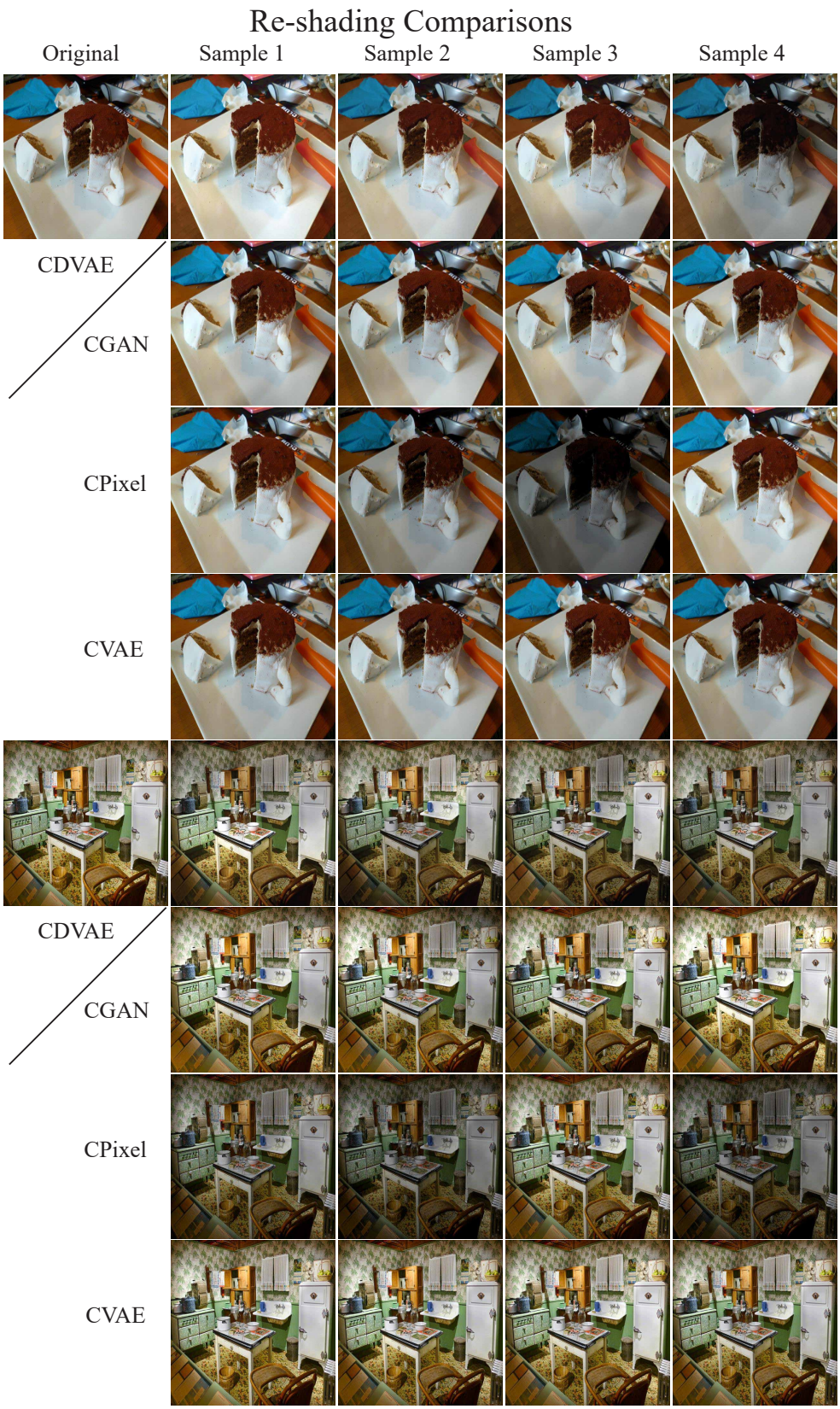

Figure 23. Photo relighting results (part 15). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

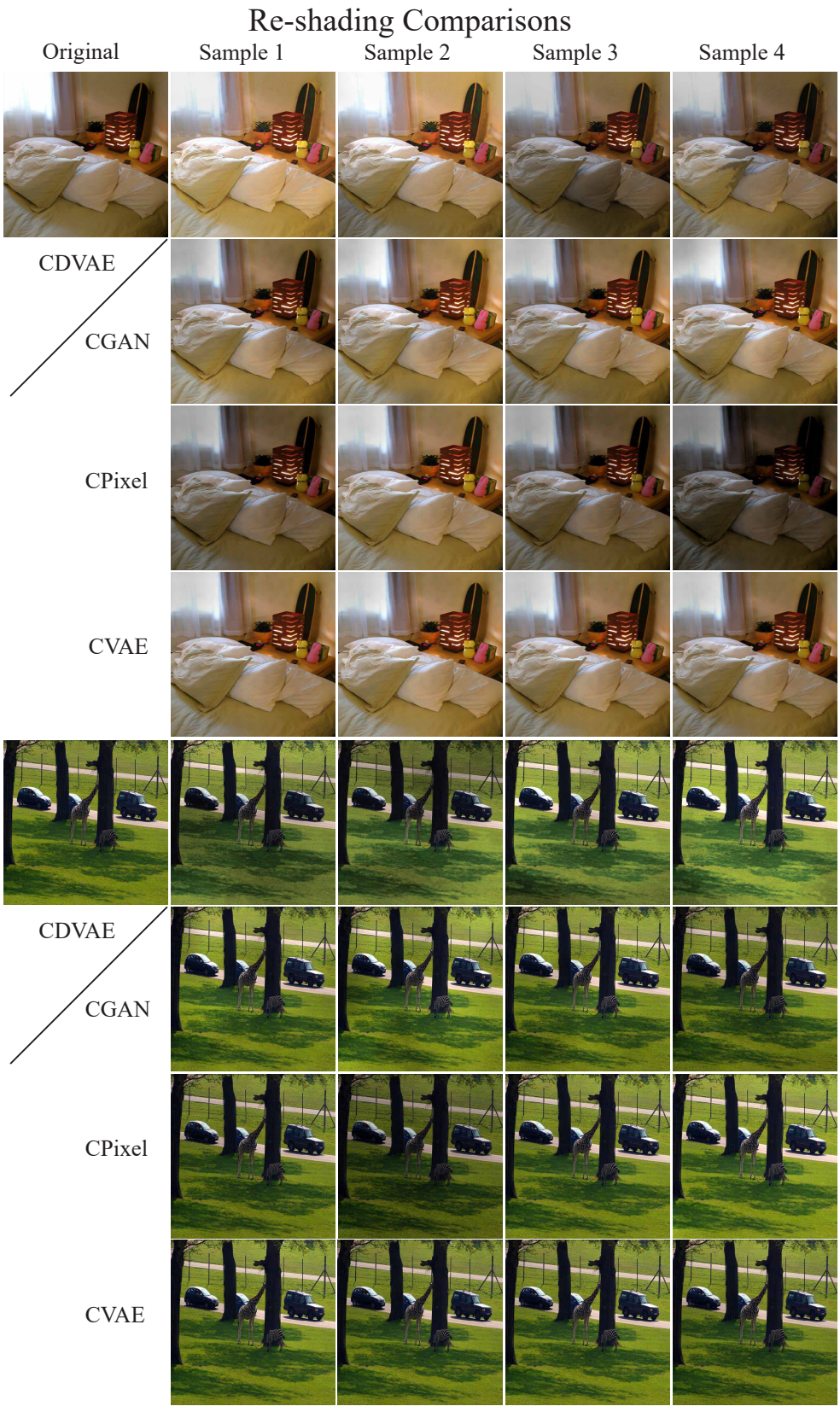

Figure 24. Photo relighting results (part 16). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

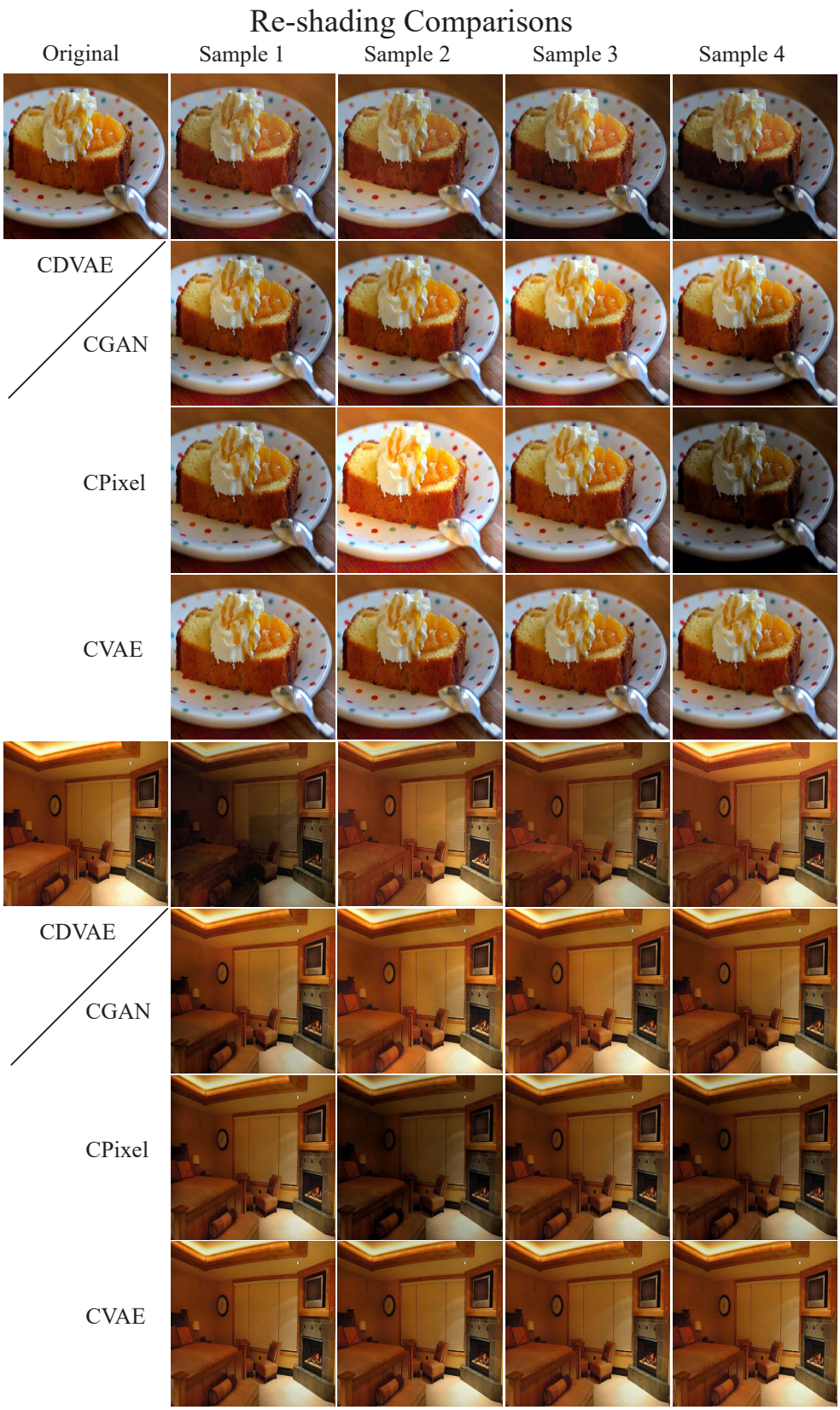

Figure 25. Photo relighting results (part 17). Photo relighting results with CGAN tend to have less variety and be less reasonable; results with CPixel tend to be extreme and random, and they also have less spatial structures; results with CVAE suffers from mode collapsion and have limited variety.

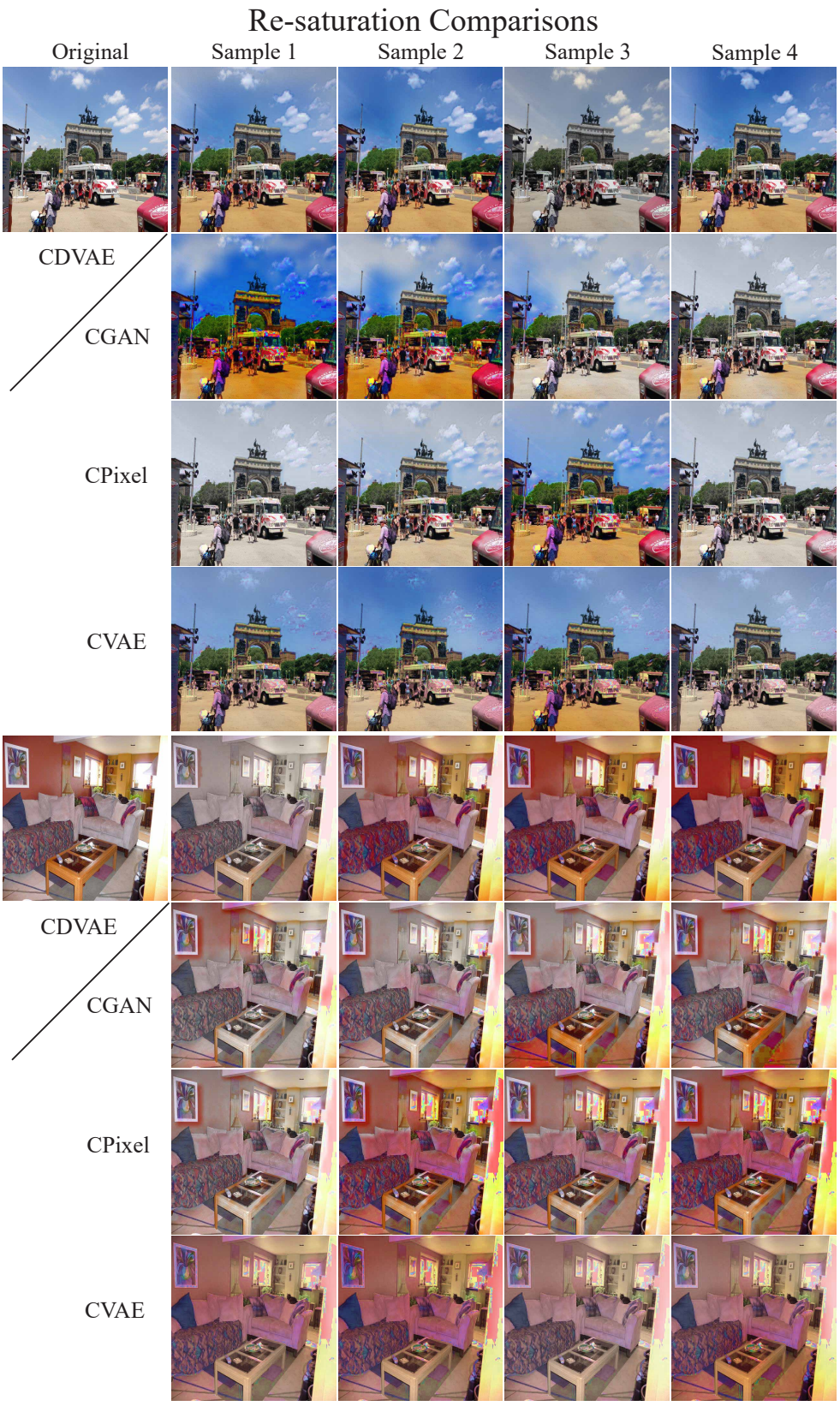

Figure 26. Image re-saturation results (part 1). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

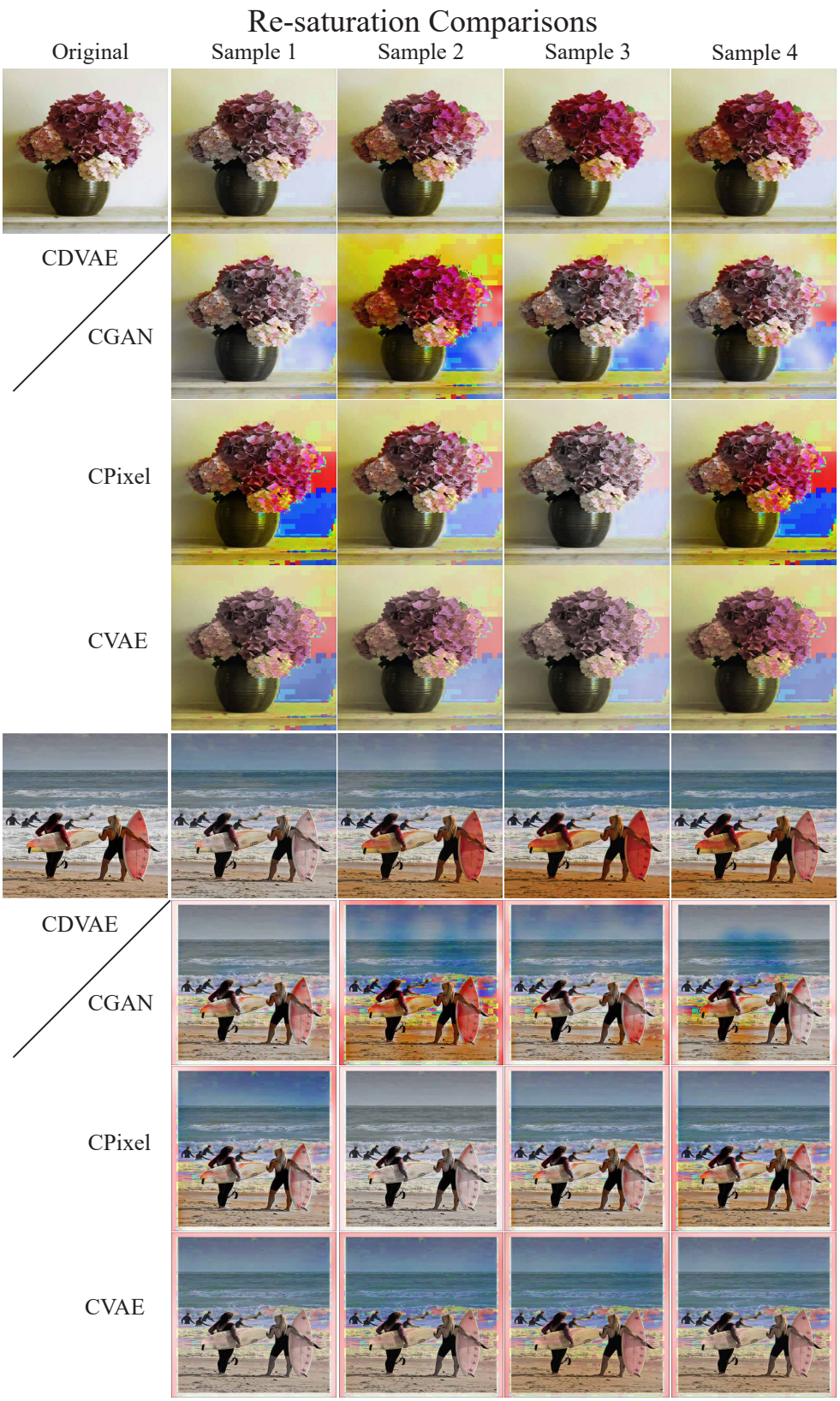

Figure 27. Image re-saturation results (part 2). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

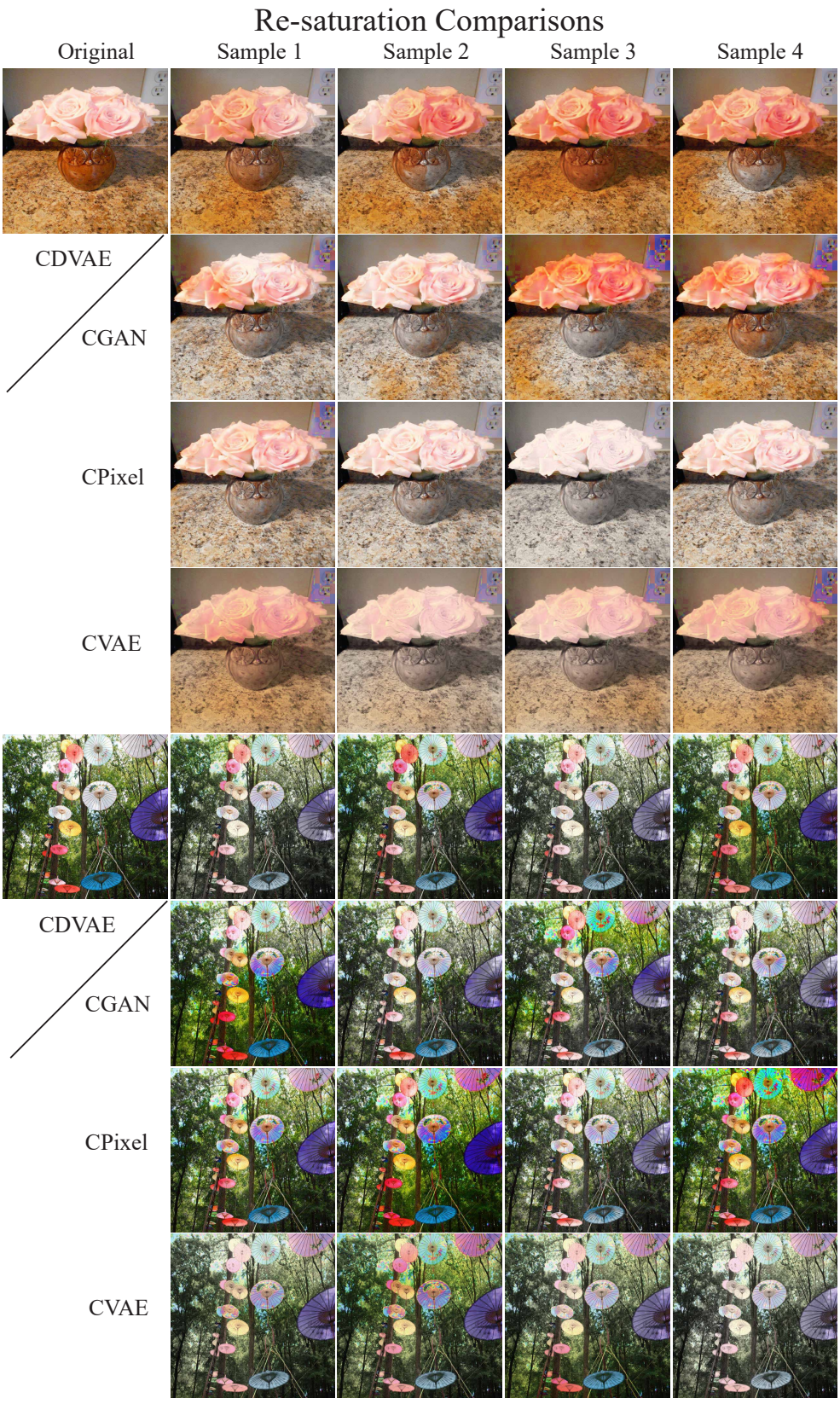

Figure 28. Image re-saturation results (part 3). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

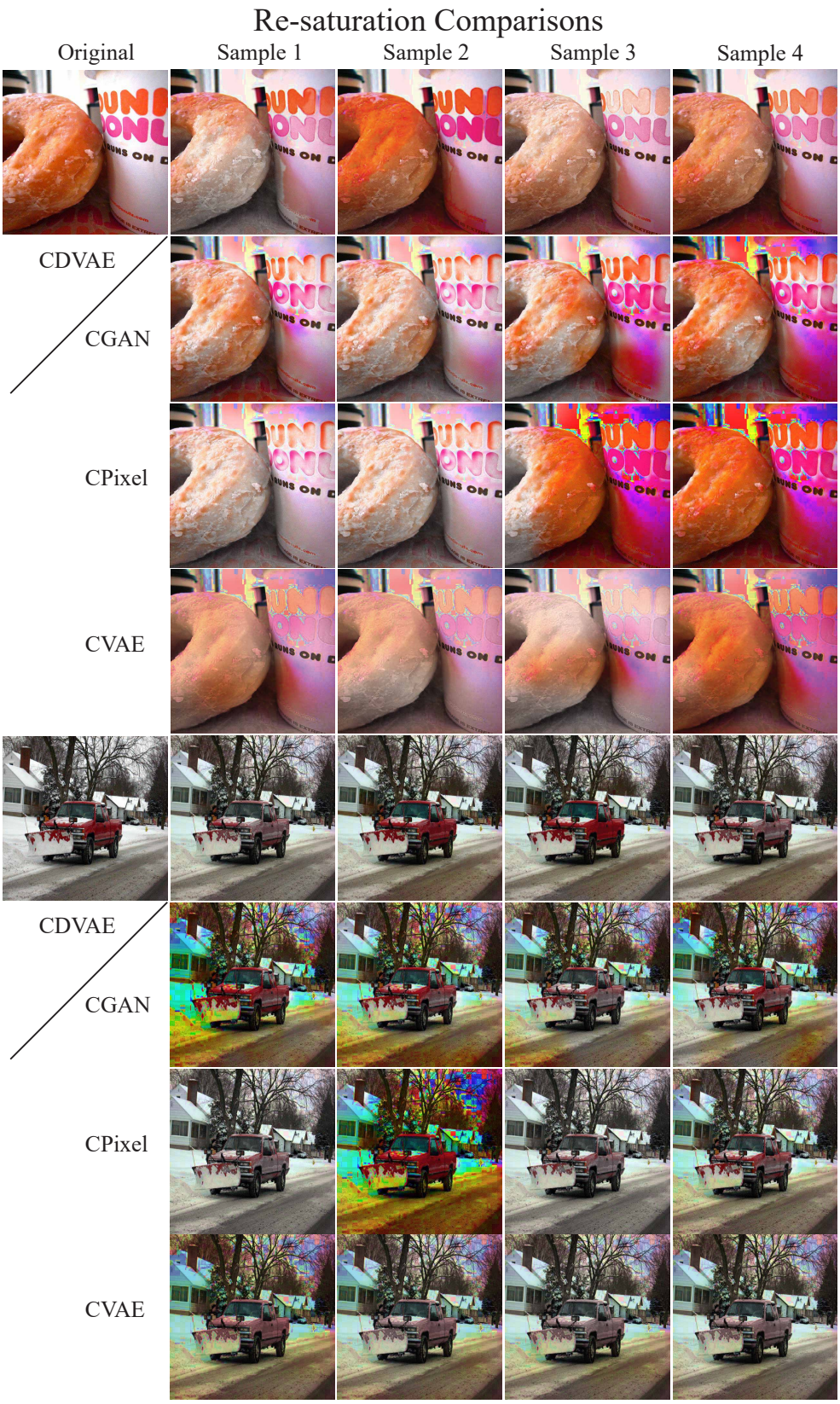

Figure 29. Image re-saturation results (part 4). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

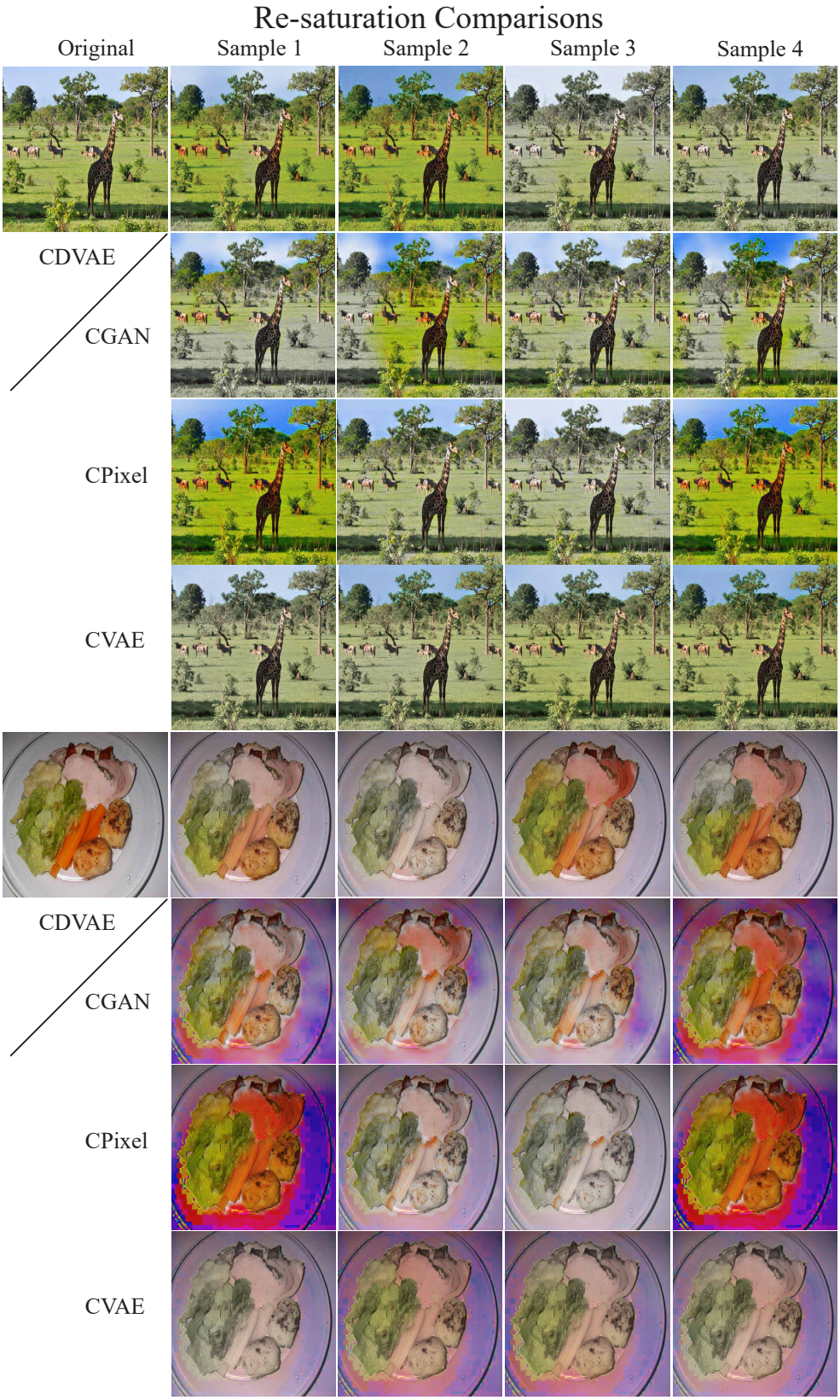

Figure 30. Image re-saturation results (part 5). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

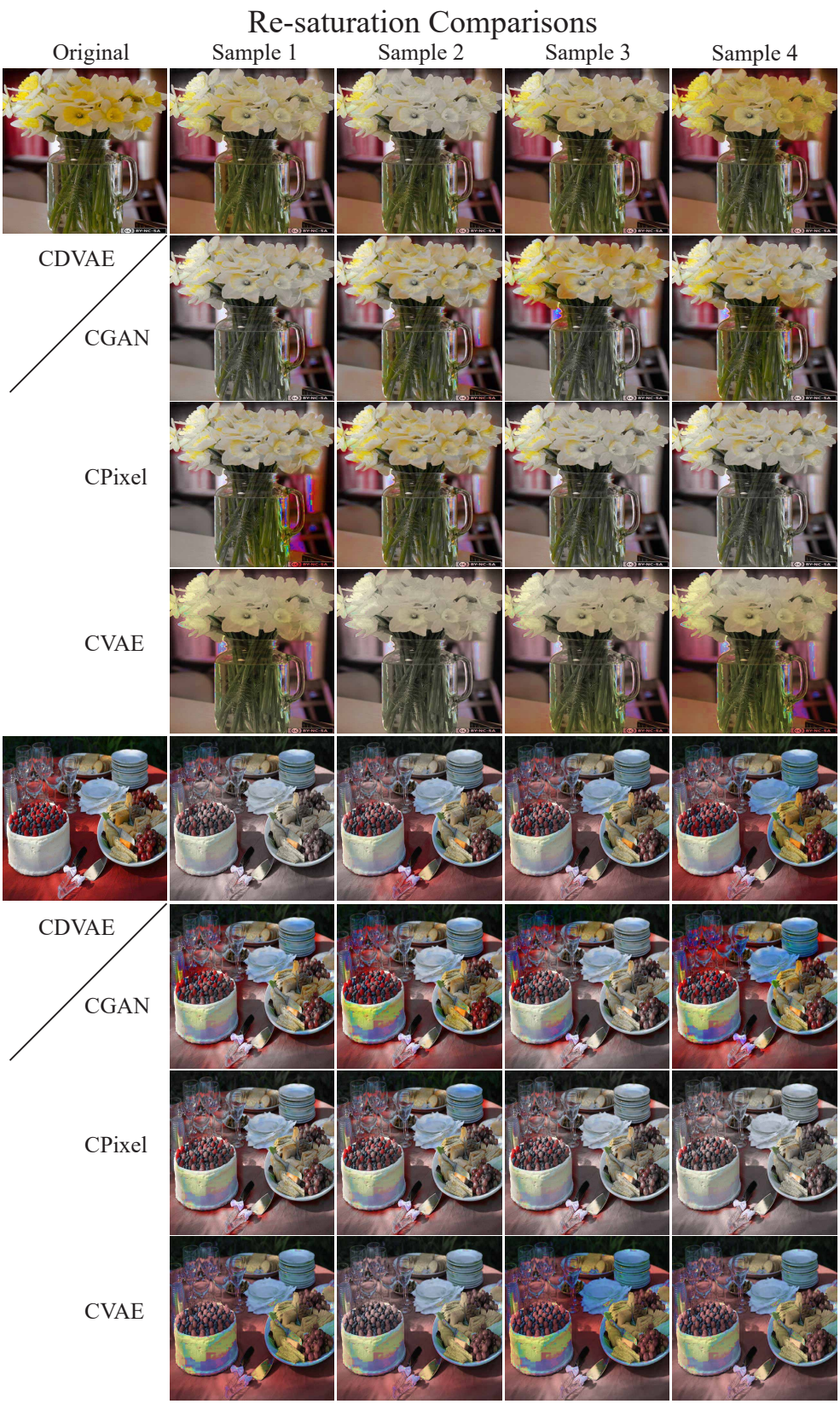

Figure 31. Image re-saturation results (part 6). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

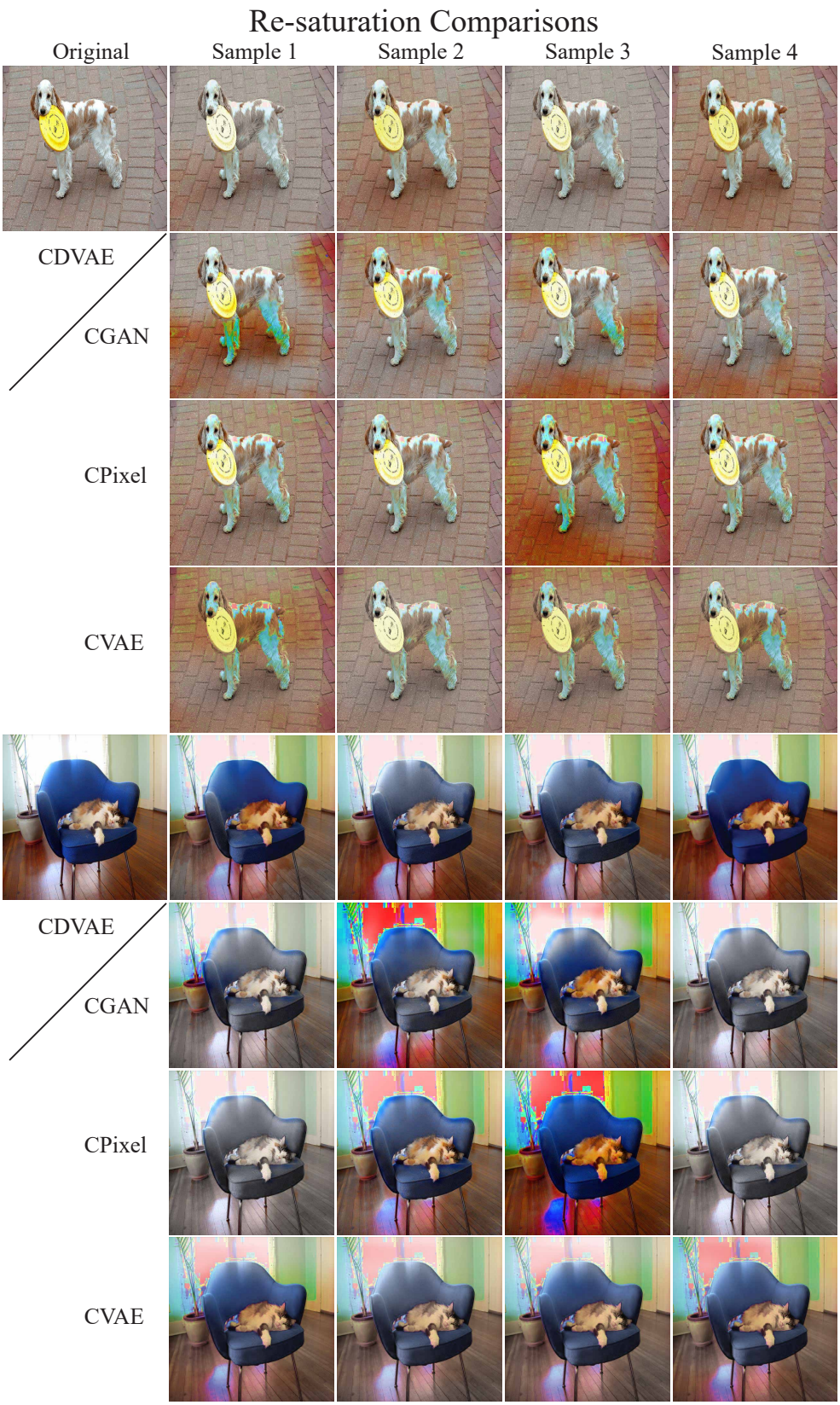

Figure 32. Image re-saturation results (part 7). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

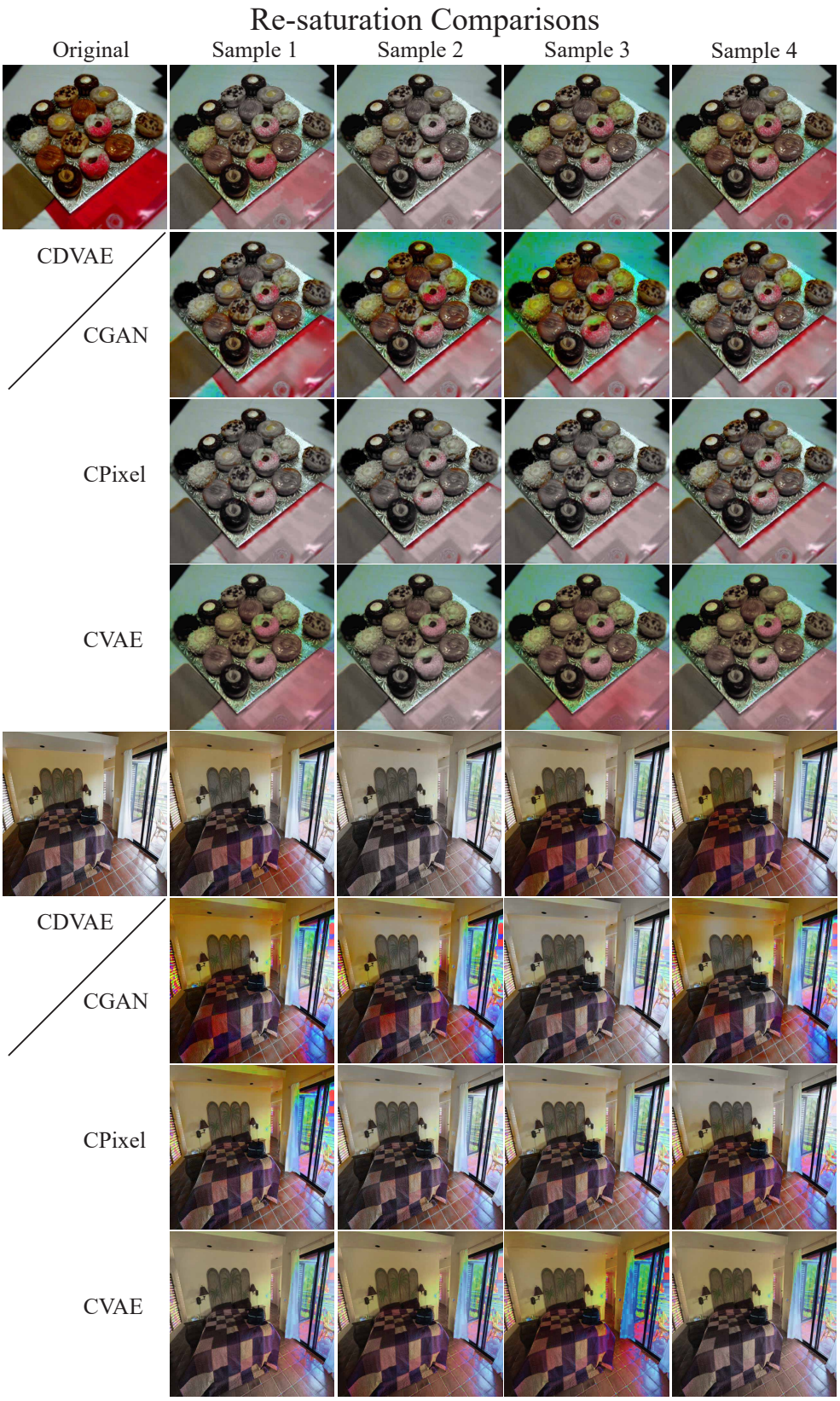

Figure 33. Image re-saturation results (part 8). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

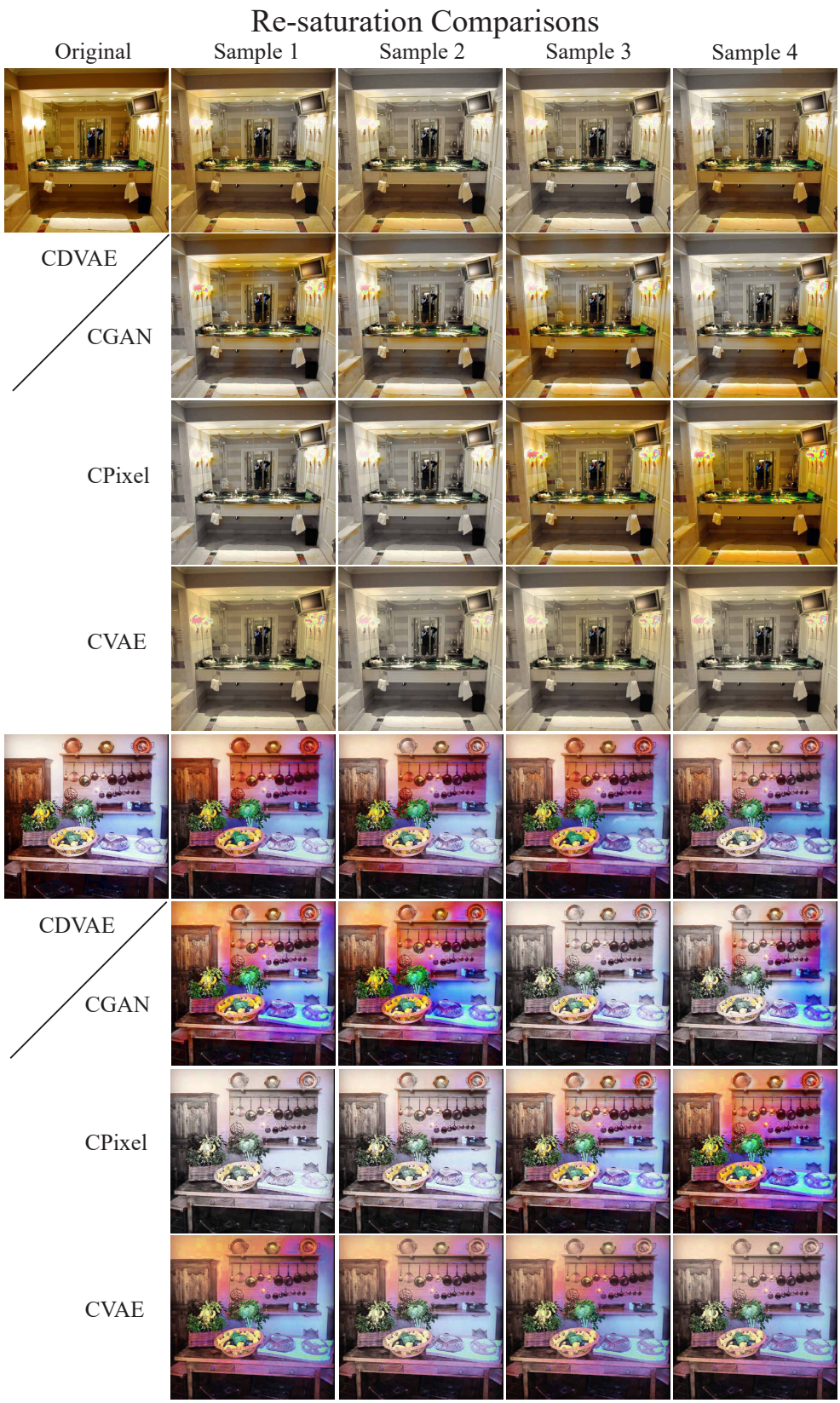

Figure 34. Image re-saturation results (part 9). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

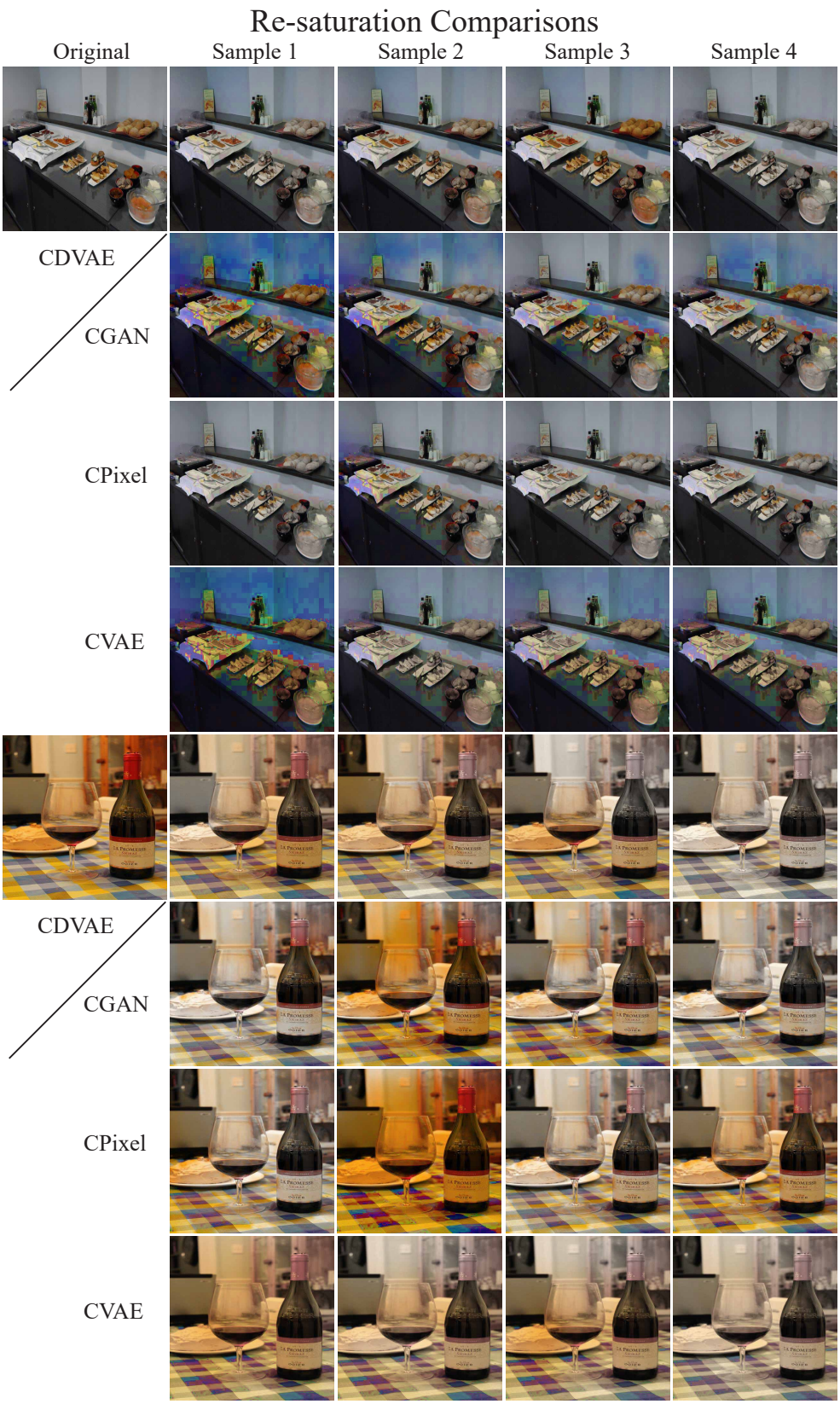

Figure 35. Image re-saturation results (part 10). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

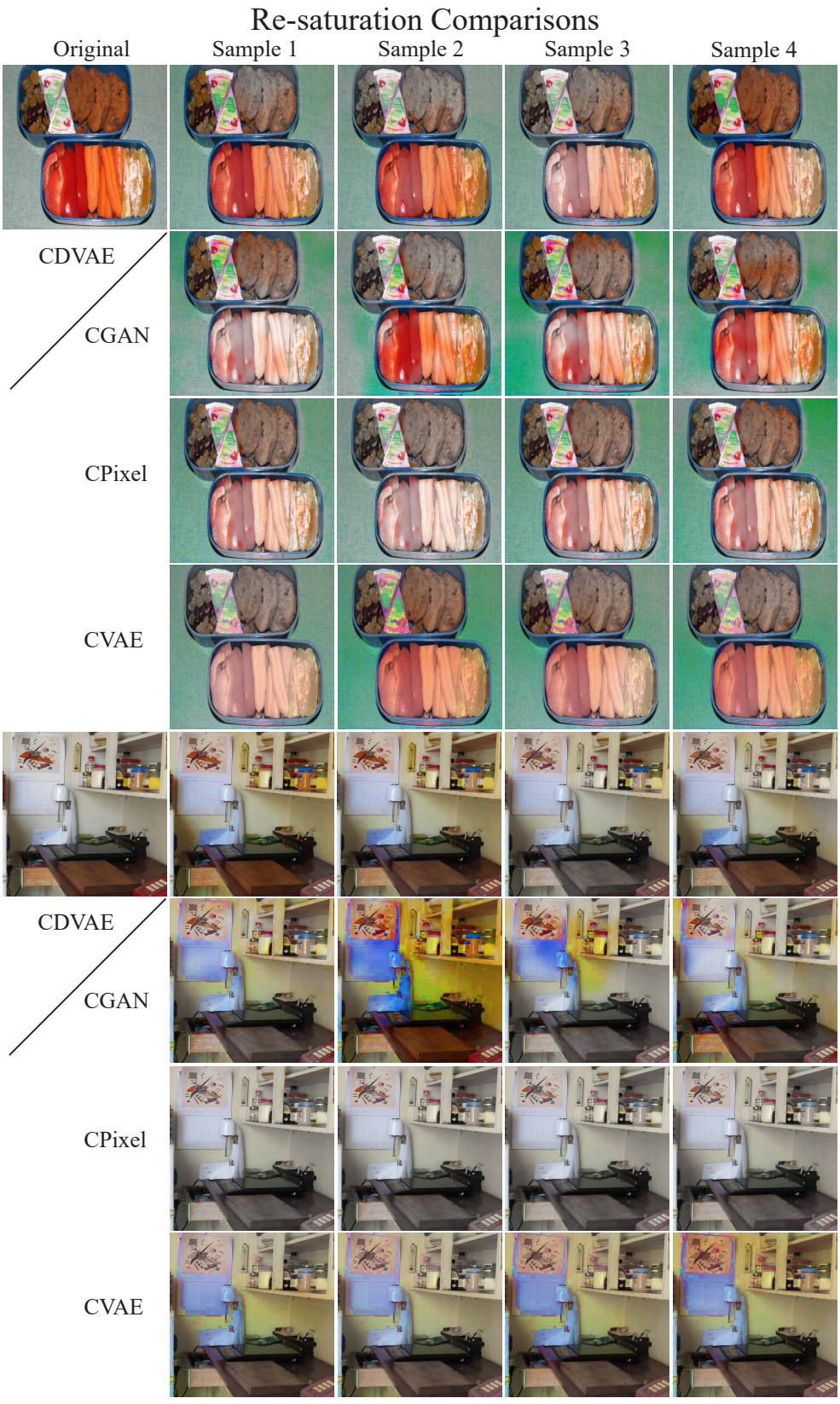

Figure 36. Image re-saturation results (part 11). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

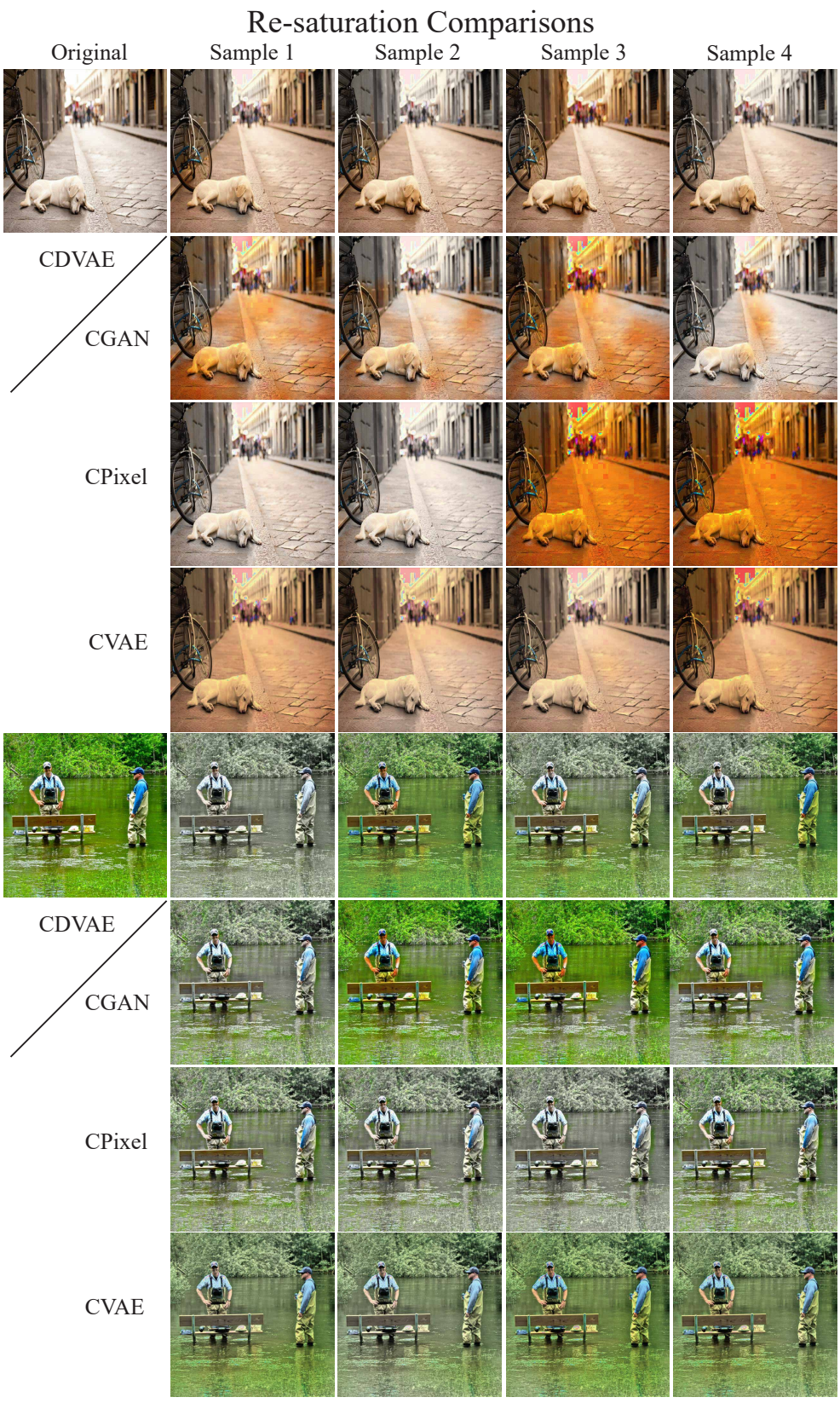

Figure 37. Image re-saturation results (part 12). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

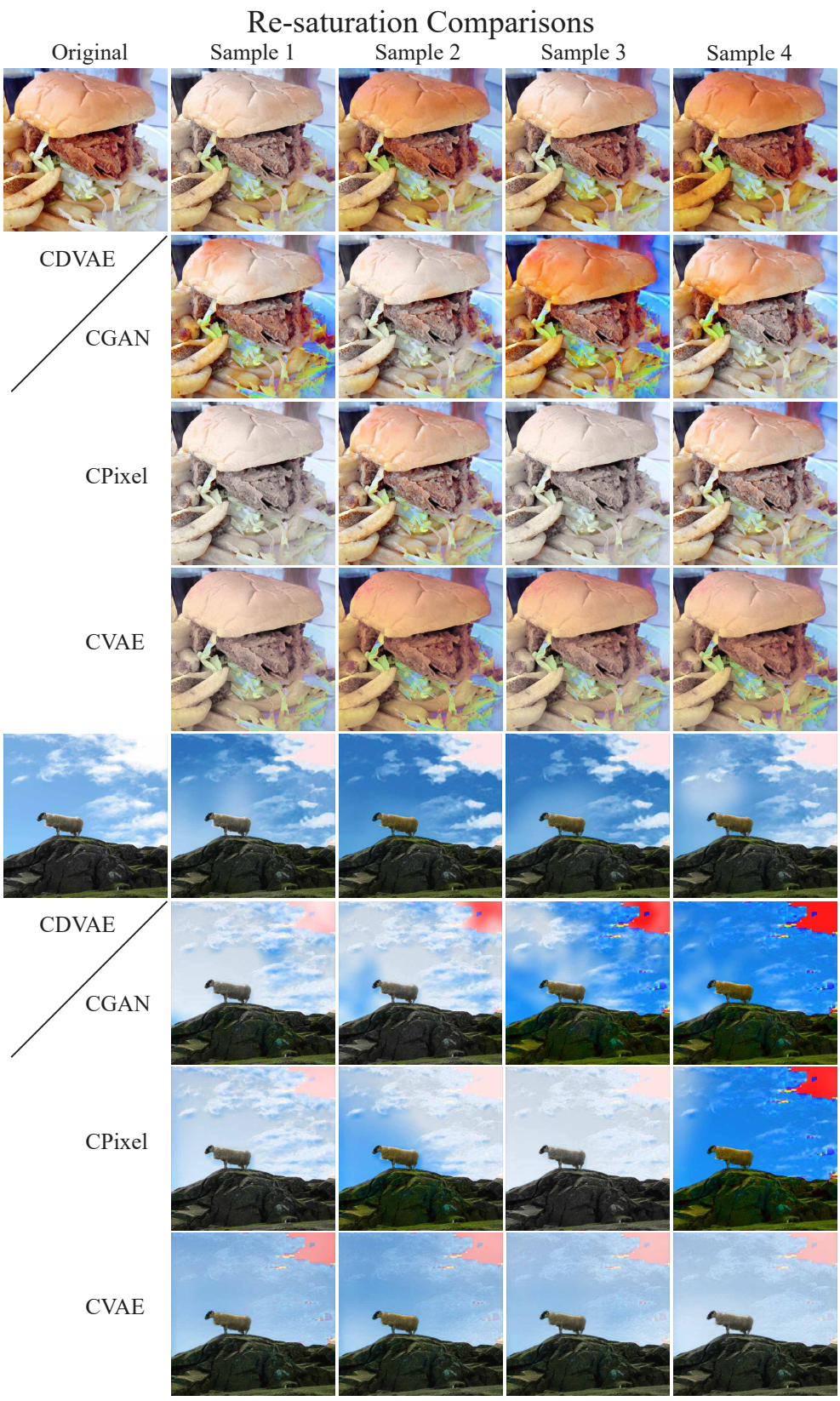

Figure 38. Image re-saturation results (part 13). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

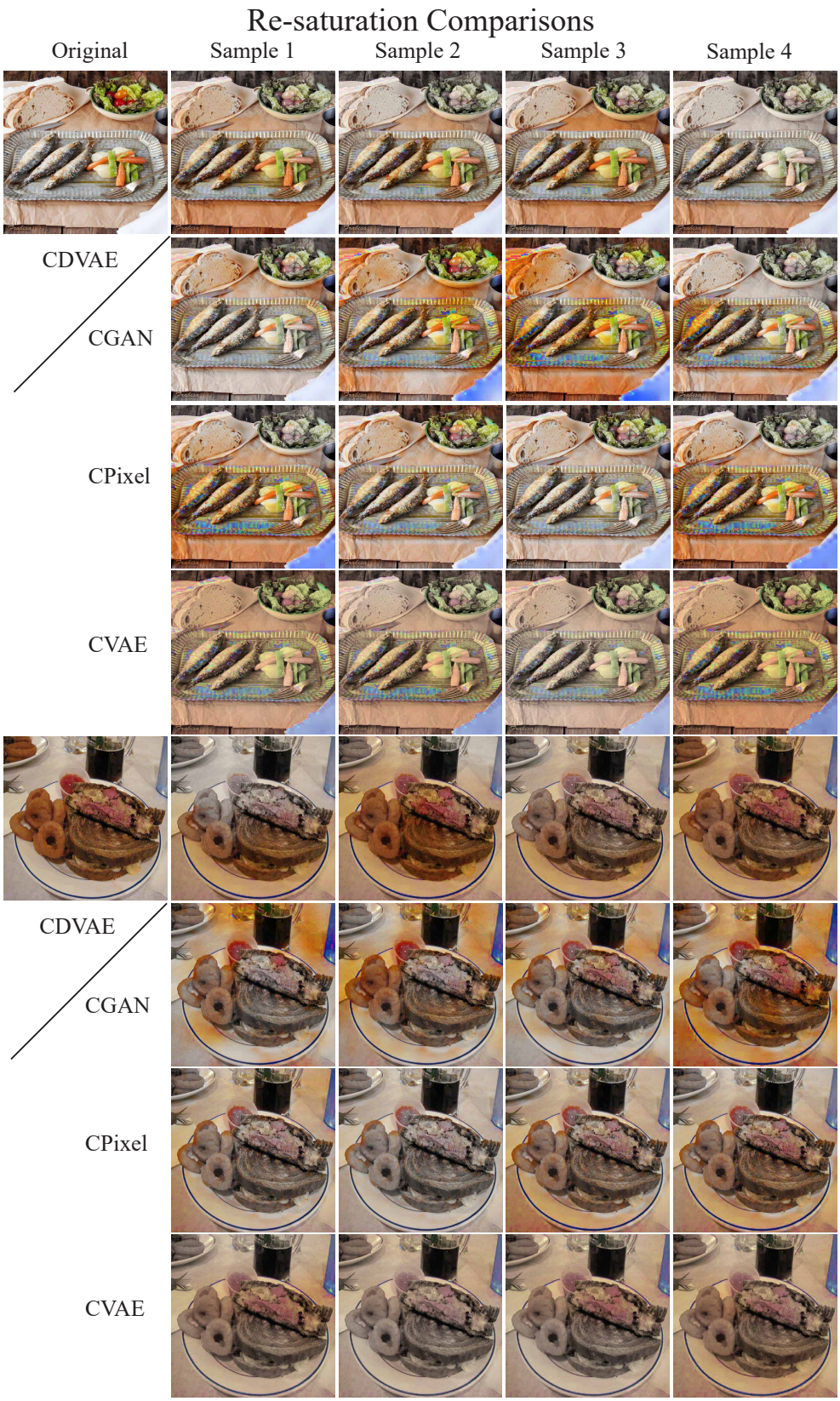

Figure 39. Image re-saturation results (part 14). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

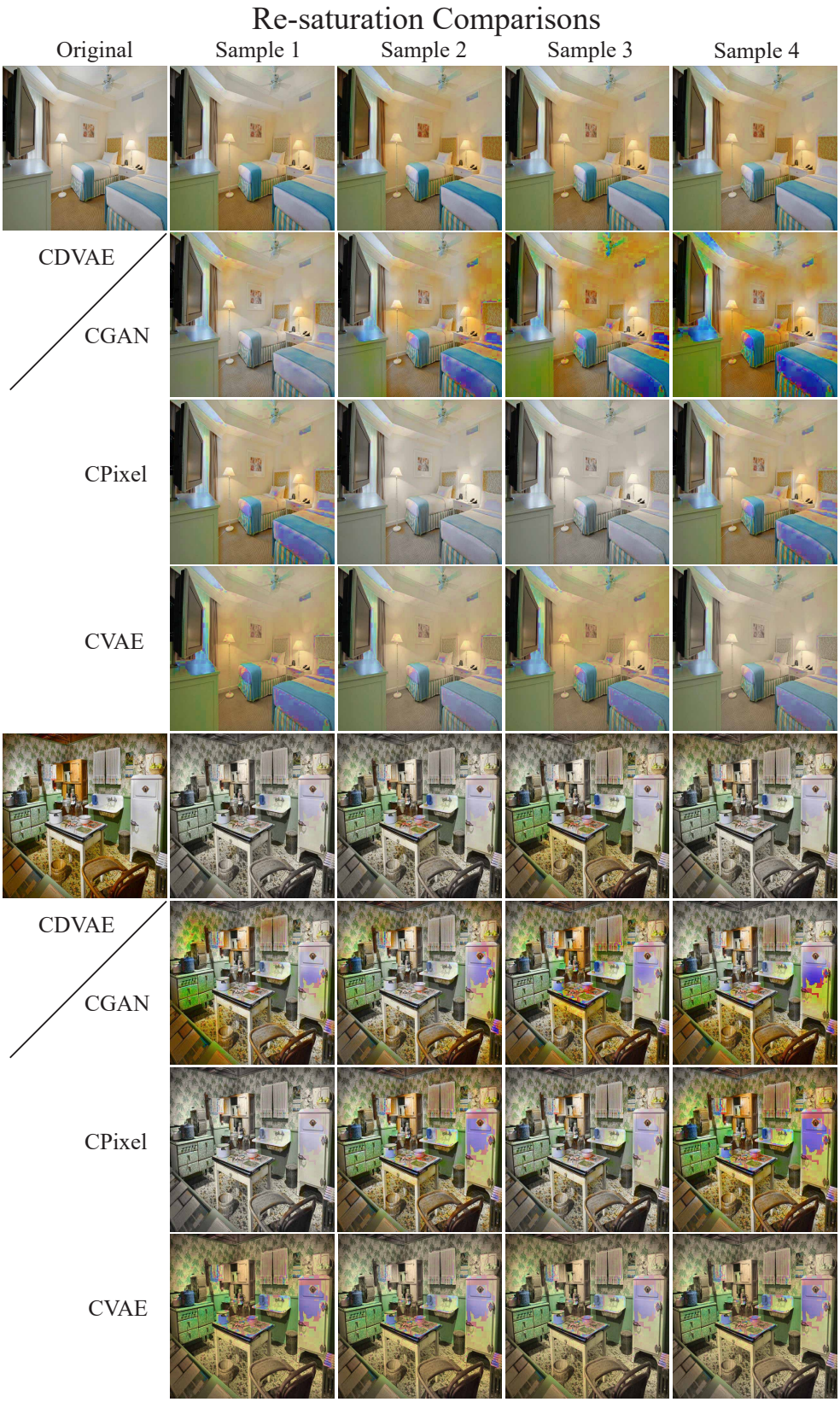

Figure 40. Image re-saturation results (part 15). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

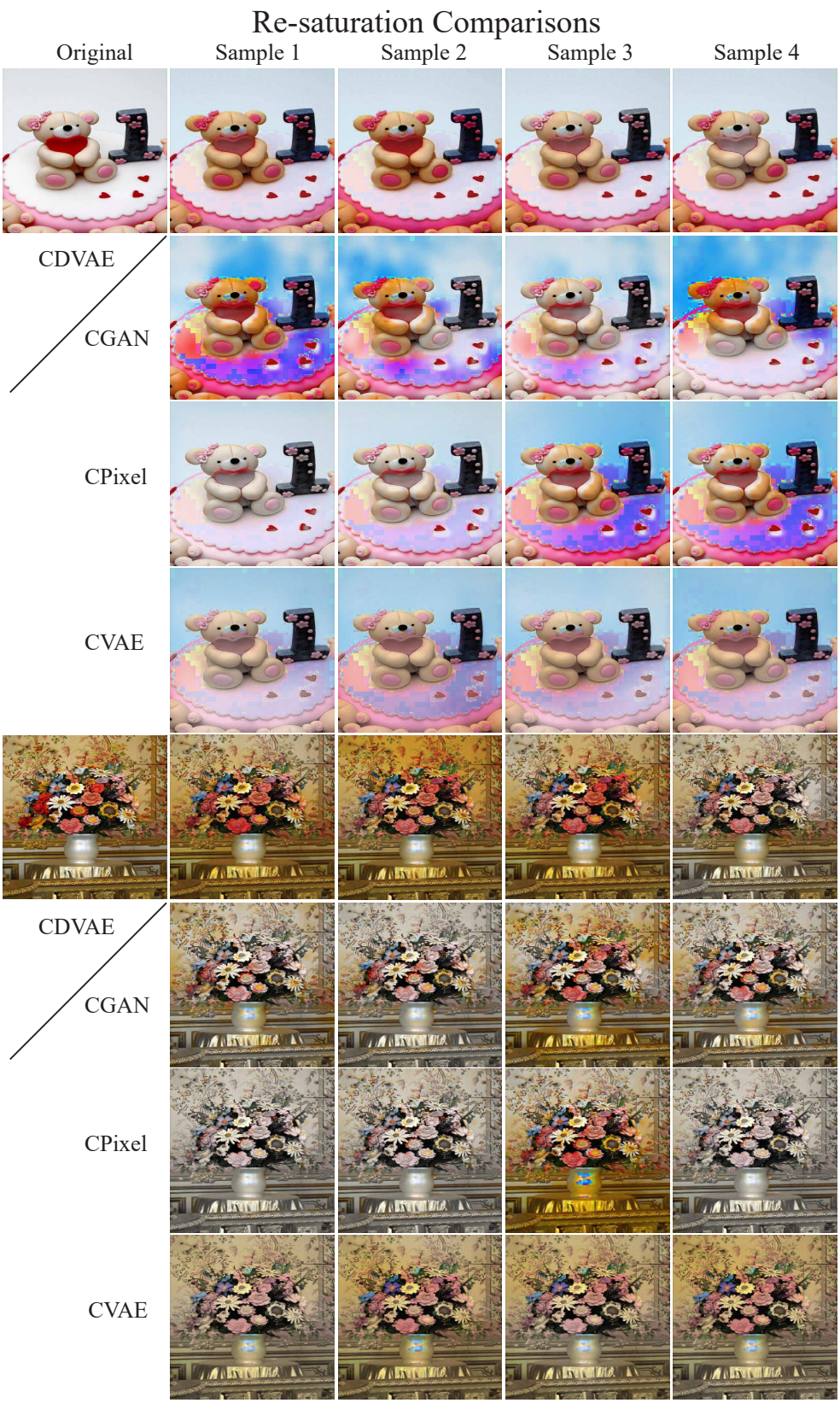

Figure 41. Image re-saturation results (part 16). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

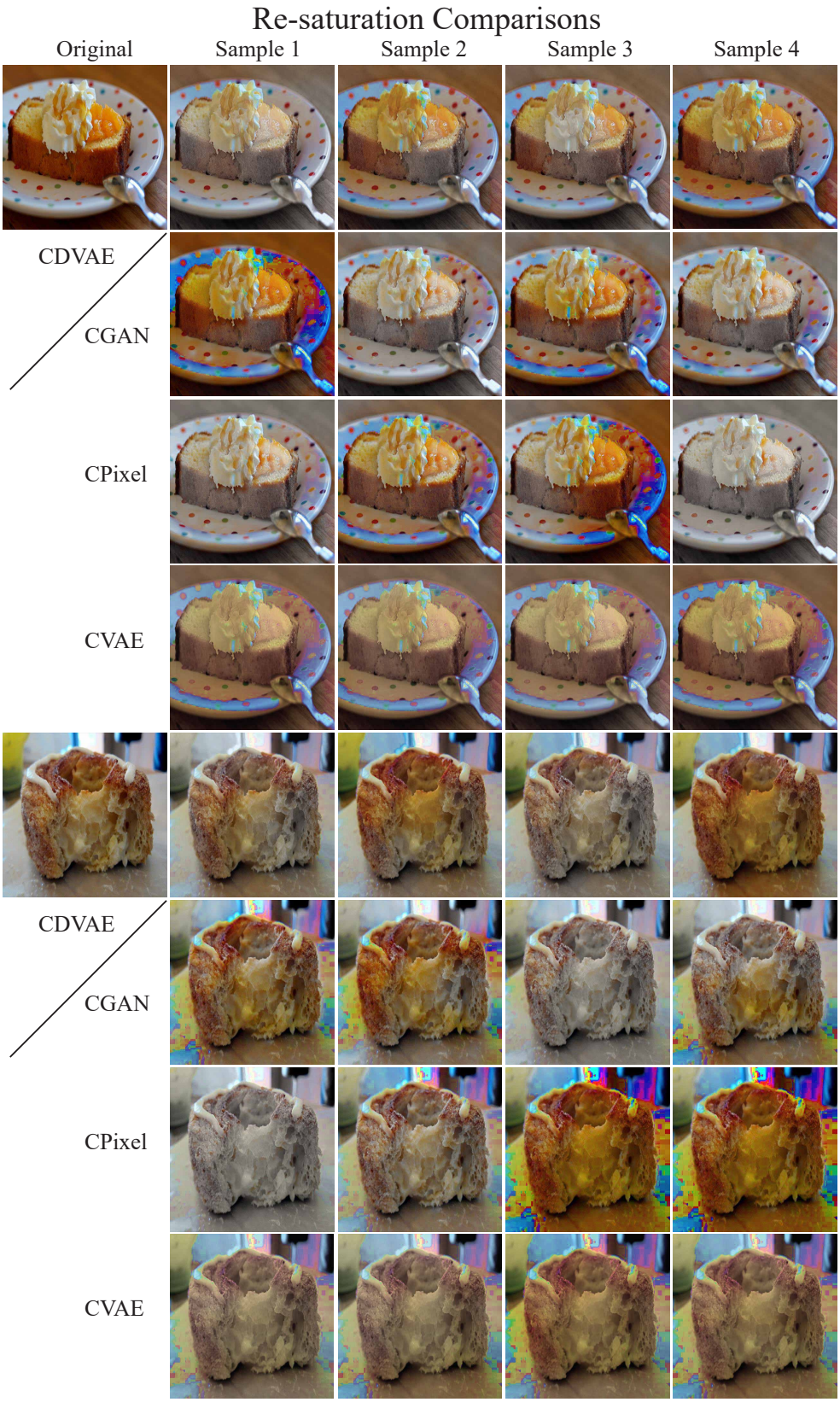

Figure 42. Image re-saturation results (part 17). Image re-saturation results with CGAN tend to ignore the image content and like random, and creates various of artifacts; results with CPixel tend to be extreme, and either like random or go into mode collapsion; results with CVAE have limited variety and creates more artifacts.

References

[1] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 1

[2] D. J. Rezende, S. Mohamed, and D. Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014. 1

[3] C. K. Sønderby, T. Raiko, L. Maaløe, S. K. Sønderby, and O. Winther. Ladder variational autoencoders. *arXiv preprint arXiv:1602.02282*, 2016. 1