Bicycle GAN

Presenter : Akilesh B

Motivation

- Existing image to image translation works produce only single output.
- This paper models a distribution of possible outputs in a conditional generative modeling setting.
- A learnt low-dimensional latent code is used in addition to input image to generate possible outputs.
- Bijective mapping between the latent encoding and output modes.

(a) Input night image

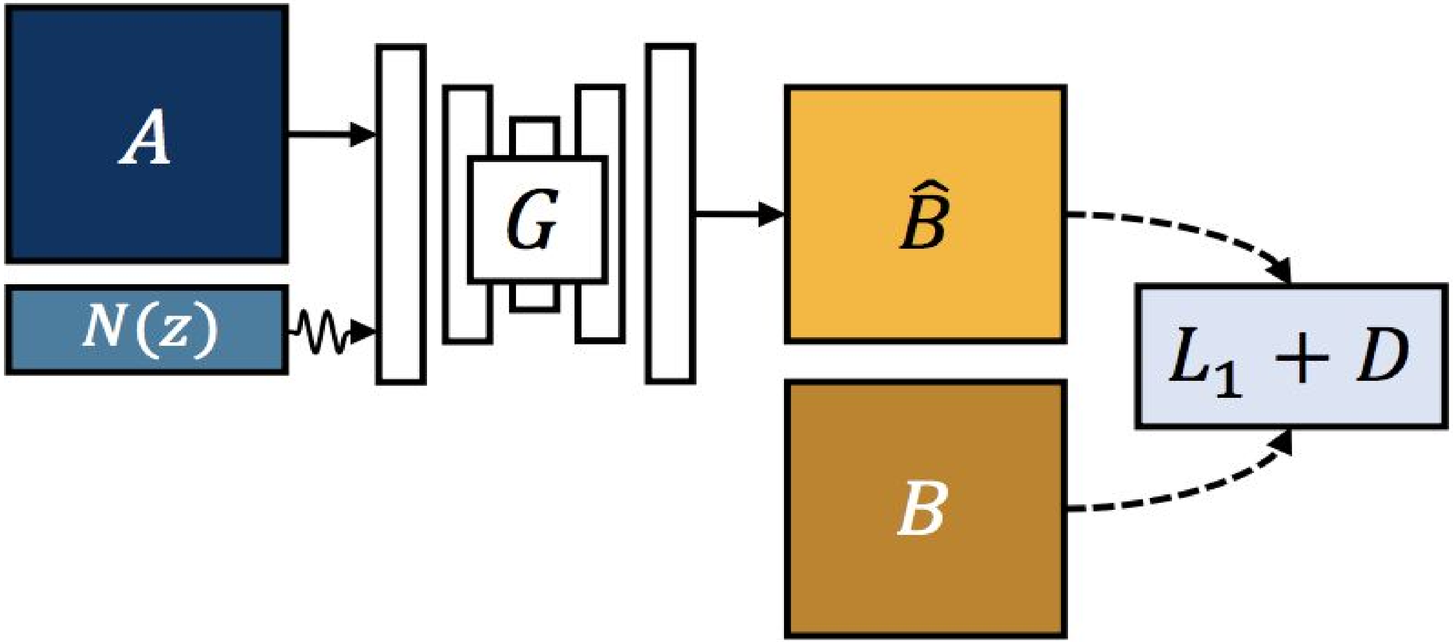
(b) Diverse day images sampled by our model

Baseline: pix2pix + noise ($z \longrightarrow \hat{B}$)

$$\mathcal{L}_{\text{GAN}}(G, D) = \mathbb{E}_{\mathbf{A},\mathbf{B}\sim p(\mathbf{A},\mathbf{B})}[\log(D(\mathbf{A}, \mathbf{B}))] + \mathbb{E}_{\mathbf{A}\sim p(\mathbf{A}),\mathbf{z}\sim p(\mathbf{z})}[\log(1 - D(\mathbf{A}, G(\mathbf{A}, \mathbf{z})))]$$

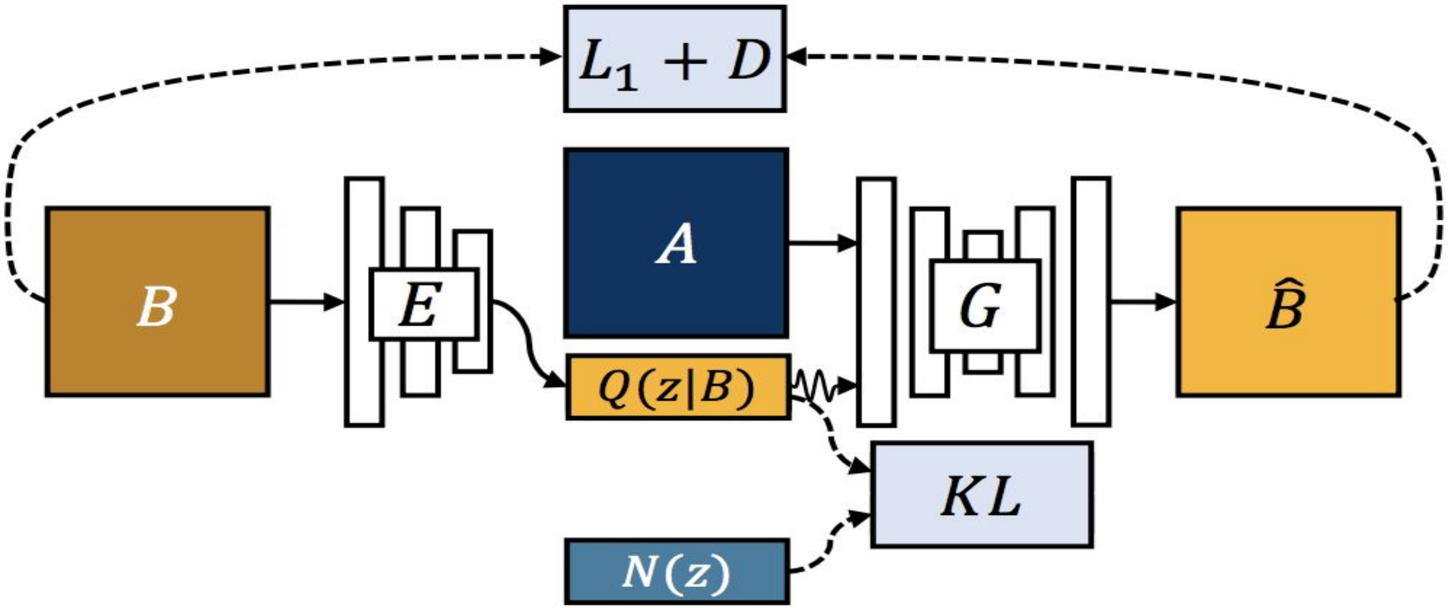
$$\mathcal{L}_1^{\text{image}}(G) = \mathbb{E}_{\mathbf{A},\mathbf{B}\sim p(\mathbf{A},\mathbf{B}),\mathbf{z}\sim p(\mathbf{z})}||\mathbf{B} - G(\mathbf{A}, \mathbf{z})||_1$$

$$G^* = \arg\min_G \max_D \quad \mathcal{L}_{\text{GAN}}(G, D) + \lambda \mathcal{L}_1^{\text{image}}(G)$$



(b) Training pix2pix+noise

cVAE-GAN ($B \longrightarrow z \longrightarrow \hat{B}$)



(c) Training cVAE-GAN

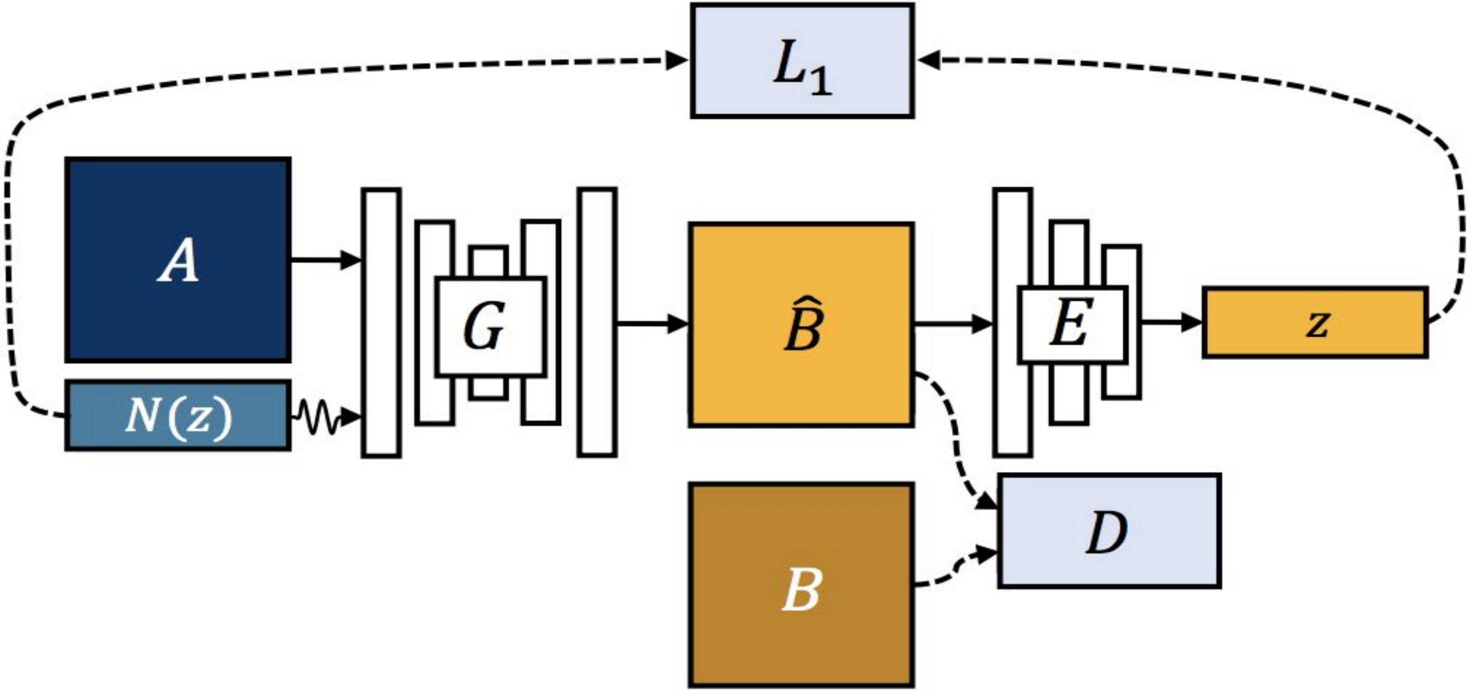
cVAE-GAN equations

$$\mathcal{L}_{\text{GAN}}^{\text{VAE}} = \mathbb{E}_{\mathbf{A},\mathbf{B}\sim p(\mathbf{A},\mathbf{B})}[\log(D(\mathbf{A},\mathbf{B}))] + \mathbb{E}_{\mathbf{A},\mathbf{B}\sim p(\mathbf{A},\mathbf{B}),\mathbf{z}\sim E(\mathbf{B})}[\log(1 - D(\mathbf{A}, G(\mathbf{A},\mathbf{z})))]$$

$$\mathcal{L}_{\text{KL}}(E) = \mathbb{E}_{\mathbf{B}\sim p(\mathbf{B})}[\mathcal{D}_{\text{KL}}(E(\mathbf{B}) || \mathcal{N}(0, I))],$$

$$G^*, E^* = \arg\min_{G,E}\max_{D} \quad \mathcal{L}_{\text{GAN}}^{\text{VAE}}(G, D, E) + \lambda\mathcal{L}_1^{\text{VAE}}(G, E) + \lambda_{\text{KL}}\mathcal{L}_{\text{KL}}(E).$$

cLR-GAN ($z \longrightarrow \hat{B} \longrightarrow \hat{z}$)



(d) Training cLR-GAN

cLR-GAN equations

$$\mathcal{L}_1^{\text{latent}}(G, E) = \mathbb{E}_{\mathbf{A} \sim p(\mathbf{A}), \mathbf{z} \sim p(\mathbf{z})} ||\mathbf{z} - E(G(\mathbf{A}, \mathbf{z}))||_1$$

$$G^*, E^* = \arg \min_{G,E} \max_{D} \quad \mathcal{L}_{\text{GAN}}(G, D) + \lambda_{\text{latent}} \mathcal{L}_1^{\text{latent}}(G, E)$$

Bicycle GAN

$$G^*, E^* = \arg\min_{G,E}\max_{D} \quad \mathcal{L}_{\text{GAN}}^{\text{VAE}}(G, D, E) + \lambda\mathcal{L}_1^{\text{VAE}}(G, E)$$

$$+\mathcal{L}_{\text{GAN}}(G, D) + \lambda_{\text{latent}}\mathcal{L}_1^{\text{latent}}(G, E) + \lambda_{\text{KL}}\mathcal{L}_{\text{KL}}(E)$$

$\boldsymbol{\lambda}_{\text{latent}} = 0.5$, $\boldsymbol{\lambda}_{\text{KL}} = 0.01$, $\boldsymbol{\lambda} = 10$