Wasserstein Auto-Encoders

2018.10.04
Ridge-i 論文よみかい

Masanari Kimura
mkimura@ridge-i.com

Abstract

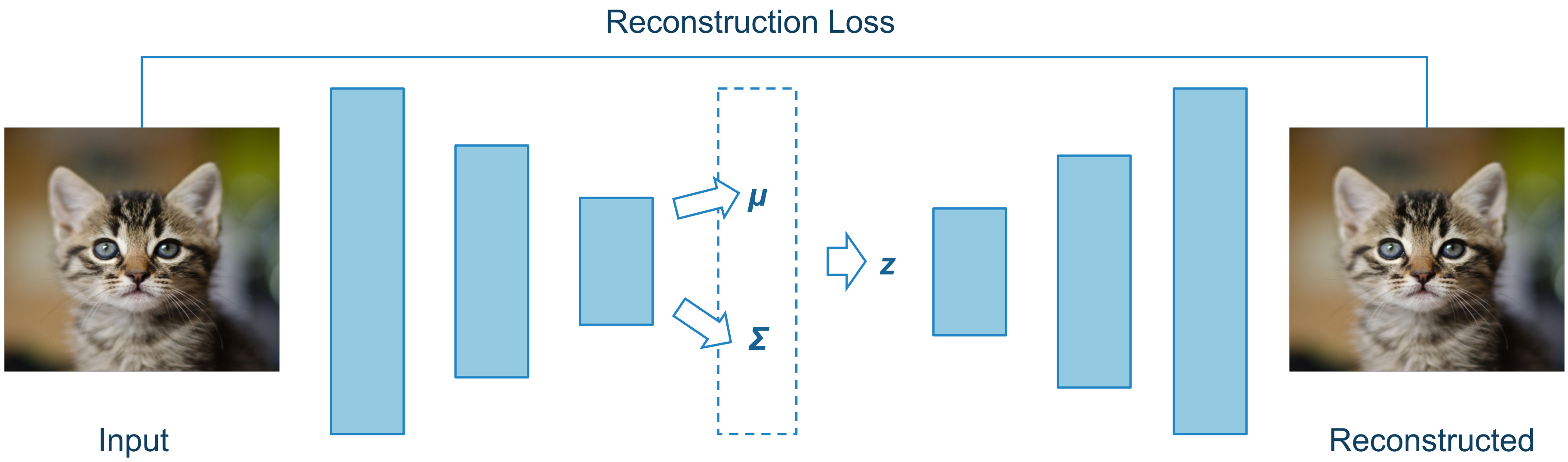
- ICLR2018採択論文 [1]
- 生成モデルの学習手法の提案.

Contributions

- Optimal transport costを最小化するように学習するWAEの提案
- GANベースとMMDベースの二つのregularizersを提案
- VAEの安定性と良好な潜在空間を保ったまま，より綺麗な画像の生成が可能

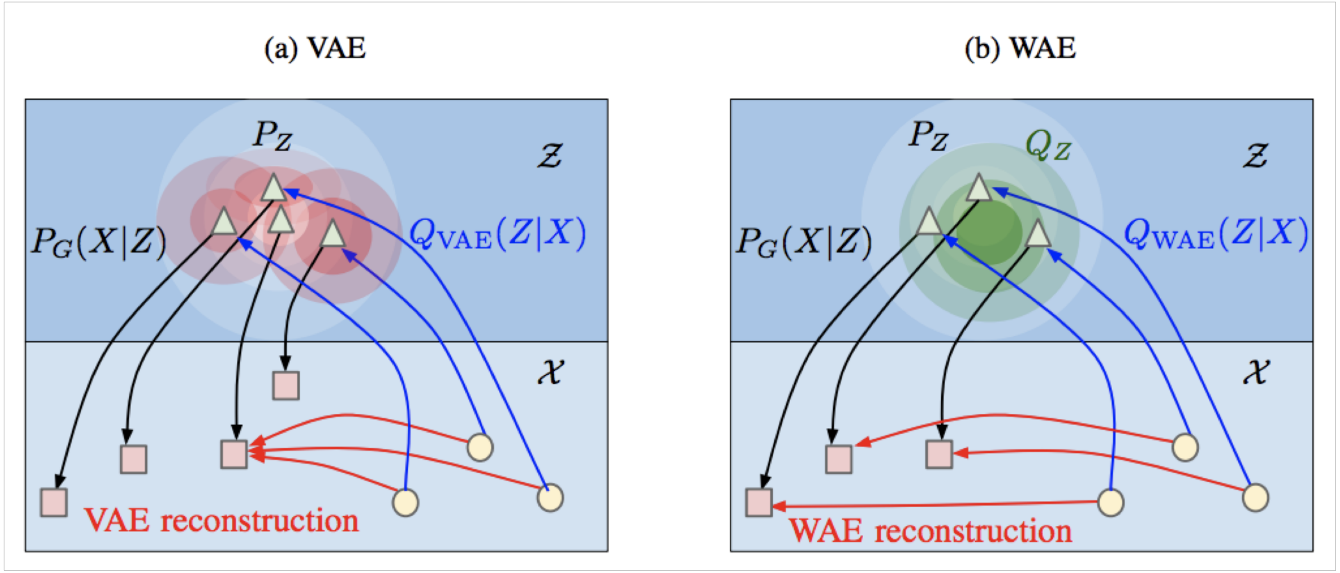
Variational AutoEncoder (VAE)

- 潜在変数zが$P_z = N(0, I_d)$からサンプリングされると仮定して学習

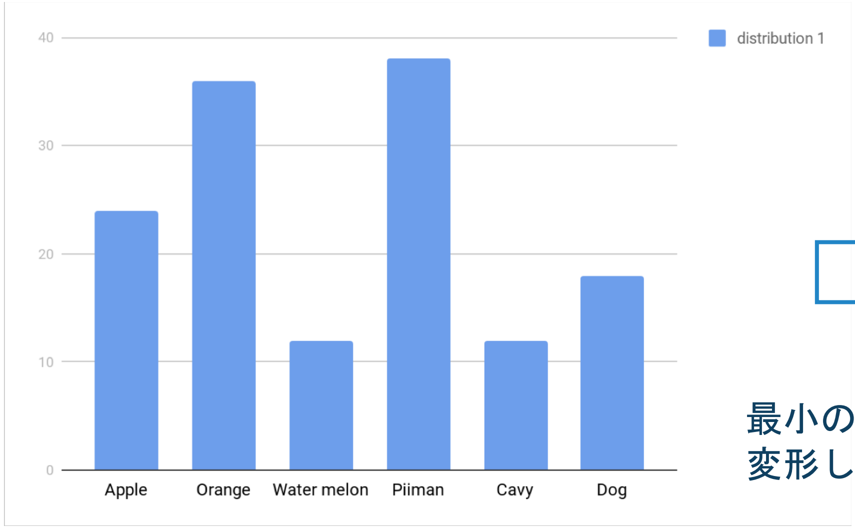


VAE vs WAE

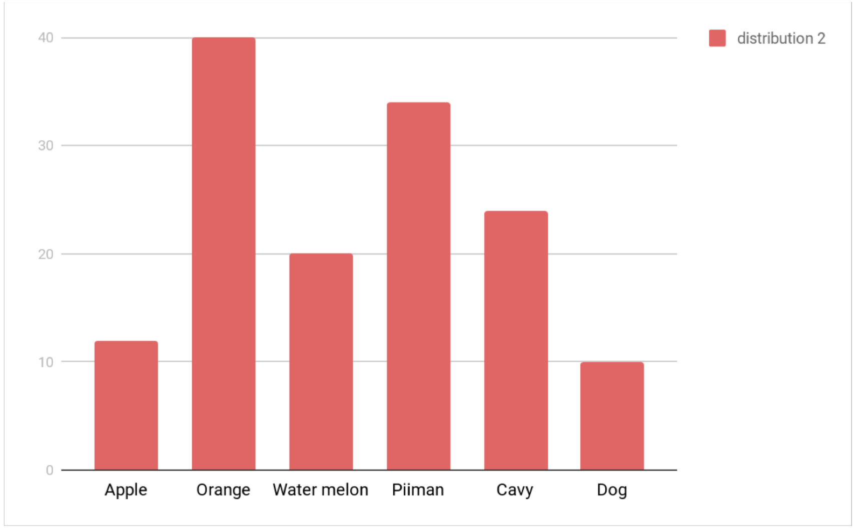
- VAE: 全てのデータ点が潜在空間上でP_zに従うように学習
  - 異なるデータ点のマージンが交差するため，デコード時に問題が起きる
- WAE: 異なるデータ点は潜在空間上で互いに離れるように学習



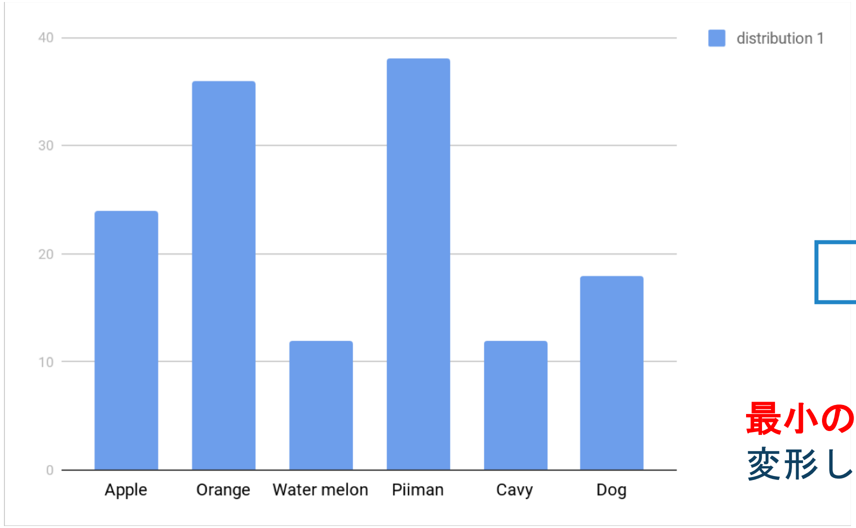
- 多くの確率分布間のダイバージェンスは，最適輸送問題によって導出される [2]



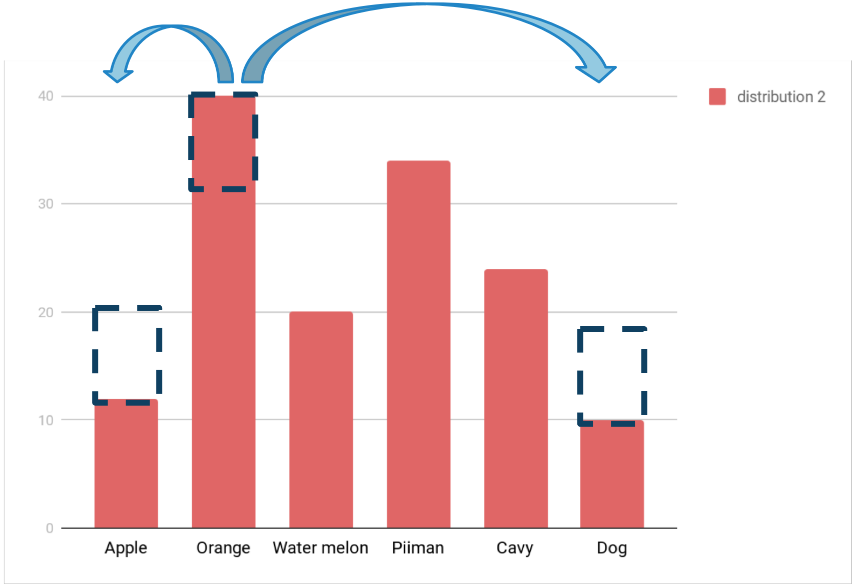
最小の労力で
変形したい



- 多くの確率分布間のダイバージェンスは，**最適輸送問題**によって導出される [2]



最小の労力で
変形したい



Orangeと間違えたものは，
Dogに運ぶよりAppleに運ぶほうが楽

Optimal Transport Problem

- 以下の問題を考える
  - $\mathrm{c(X, Y)}$ は任意のコスト関数
  - $\mathrm{P}(\mathrm{X} \sim P_X, Y \sim P_G)$はrealとgeneratedの結合分布から得られる$\mathrm{(X, Y)}$の集合

$$W_c(P_X, P_G) := \inf_{\Gamma \in \mathcal{P}(X \sim P_X, Y \sim P_G)} \mathbb{E}_{(X,Y)\sim\Gamma}[c(X,Y)],$$

- 特に$\mathrm{c}(\mathrm{X}, \mathrm{Y}) = \mathrm{d}^p(X, Y)$の時, $\mathrm{W}_c(X, Y)$のp乗根 $\mathrm{W}_p(X, Y)$ をp-Wasserstein distanceという

$$W_c(P_X, P_G) := \inf_{\Gamma \in \mathcal{P}(X \sim P_X, Y \sim P_G)} \mathbb{E}_{(X,Y)\sim\Gamma}[c(X, Y)],$$

Ridge-i

- ここで, $\mathrm{c(x, y) = d(x, y)}$，つまりp=1の時，次の双対問題が成り立つ：
  - Kantorovich-Rubinstein duality

$$W_1(P_X, P_G) = \underbrace{\sup_{f \in \mathcal{F}_L}}_{\text{1-Lipschitz関数}} \mathbb{E}_{X \sim P_X}[f(X)] - \mathbb{E}_{Y \sim P_G}[f(Y)],$$

Latent Variable Model

- 潜在表現zがある分布からサンプリングされているとする
- zに何らかの変形を施すことで画像空間にマッピングする時，生成モデルの確率密度は

$$p_G(x) := \int_{\mathcal{Z}} p_G(x|z) p_z(z) dz, \quad \forall x \in \mathcal{X},$$

- Auto-Encodersの枠組みでは，zはエンコーダの出力

WAE objective

Ridge-i

- 以下の結合分布について考える

$$\gamma(x,y) = \int p(y|z)q(z|x)p_X(x)dz$$

- xについて積分し，周辺化する

$$\int \gamma(x,y)dx = p_G(y) = \int\int p(y|z)q(z|x)p_X(x)dxdz$$

- 上式の，xに絡む項のみに注目すると，

$$\int q(z|x)p_X(x)dx = p_Z(z)$$

- $x \sim p_X(x), y \sim p_G(y)$ 間の最適輸送問題を考える．前スライドの期待値の部分を積分に変形して，

$$W(P_X, P_Y) = \inf_\gamma \int c(x,y)\gamma(x,y)dxdy$$

$$W(P_X, P_Y) = \inf_\gamma \int c(x,y)p(y|z)q(z|x)p_X(x)dzdxdy$$

- ここで，デコーダが決定的であると仮定すると，$p(y|z) = \delta(y - G(z))$，$\int \delta(x-a)f(x)dx = f(a)$ から

$$W(P_X, P_Z) = \int c(x, G(z))q(z|x)p_X(x)dxdz$$

$$W(P_X, P_Z) = \inf_{q(z|x)} E_X E_{q(z|x)} c(x, G(z))$$

(制約として $\int q(z|x)p_x(x)dx = p_z(z)$)

Penalized objective

- GAN-based
  or
  MMD-based

ALGORITHM 1 Wasserstein Auto-Encoder with GAN-based penalty (WAE-GAN).

Require: Regularization coefficient $\lambda > 0$.
Initialize the parameters of the encoder Q_ϕ, decoder G_θ, and latent discriminator D_γ.
while (ϕ, θ) not converged **do**
Sample $\{x_1, \dots, x_n\}$ from the training set
Sample $\{z_1, \dots, z_n\}$ from the prior P_Z
Sample $\tilde{z}_i$ from $Q_\phi(Z|x_i)$ for $i = 1, \dots, n$
Update D_γ by ascending:

$$\frac{\lambda}{n} \sum_{i=1}^{n} \log D_\gamma(z_i) + \log\bigl(1 - D_\gamma(\tilde{z}_i)\bigr)$$

Update Q_ϕ and G_θ by descending:

$$\frac{1}{n} \sum_{i=1}^{n} c\bigl(x_i, G_\theta(\tilde{z}_i)\bigr) - \lambda \cdot \log D_\gamma(\tilde{z}_i)$$

end while

ALGORITHM 2 Wasserstein Auto-Encoder with MMD-based penalty (WAE-MMD).

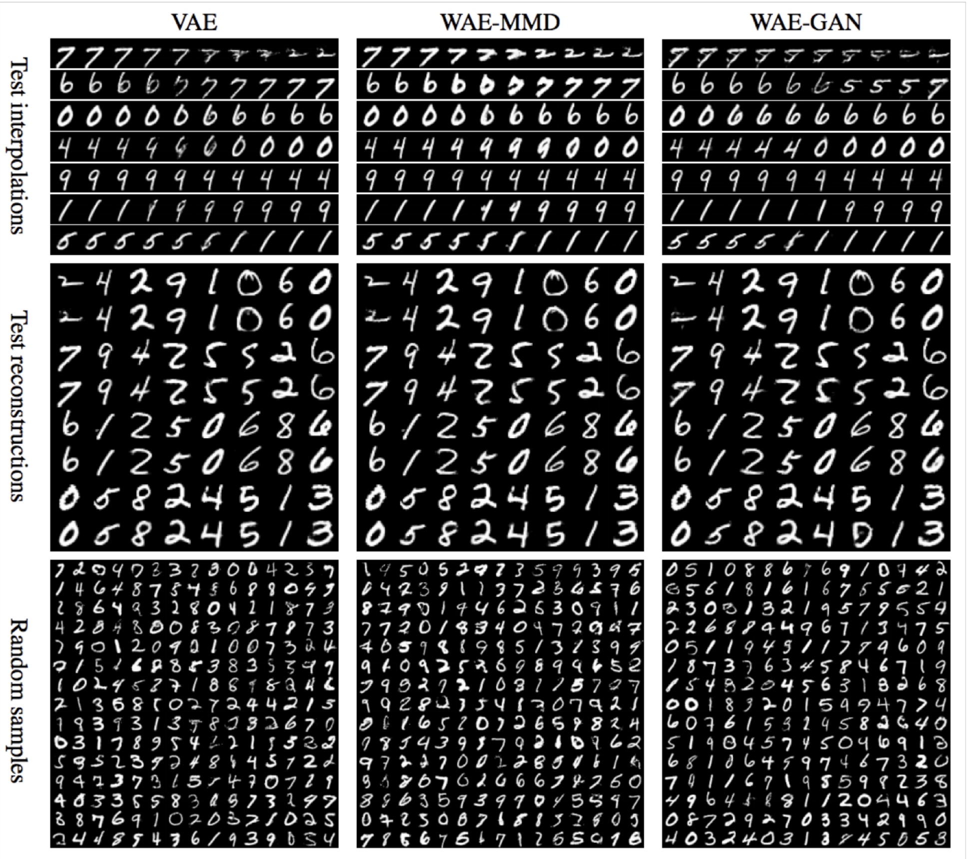
Require: Regularization coefficient $\lambda > 0$, characteristic positive-definite kernel k.
Initialize the parameters of the encoder Q_ϕ, decoder G_θ, and latent discriminator D_γ.
while (ϕ, θ) not converged **do**
Sample $\{x_1, \dots, x_n\}$ from the training set
Sample $\{z_1, \dots, z_n\}$ from the prior P_Z
Sample $\tilde{z}_i$ from $Q_\phi(Z|x_i)$ for $i = 1, \dots, n$
Update Q_ϕ and G_θ by descending:

$$\frac{1}{n} \sum_{i=1}^{n} c\bigl(x_i, G_\theta(\tilde{z}_i)\bigr) + \frac{\lambda}{n(n-1)} \sum_{\ell \neq j} k(z_\ell, z_j) + \frac{\lambda}{n(n-1)} \sum_{\ell \neq j} k(\tilde{z}_\ell, \tilde{z}_j) - \frac{2\lambda}{n^2} \sum_{\ell, j} k(z_\ell, \tilde{z}_j)$$

end while

Experimental Results

- MNISTデータセット



Ridge-i

- CelebAデータセット

Conclusions

- Optimal transport costを用いた新しい生成モデルの学習手法を提案
- 既存のVAEと比べて，より綺麗な画像の生成が可能になった

References

- [1] Ilya Tolstikhin, Olivier Bousquet, Sylvain Gelly, Bernhard Schoelkopf. "Wasserstein Auto-Encoders", International Conference on Learning Representations (ICLR). 2018.
- [2] C. Villani. Topics in Optimal Transportation. AMS Graduate Studies in Mathematics, 2003.